A Heuristic Algorithm for the Resource Assignment Problem in Satellite Telecommunication Networks

David Fernandez[1] , Sergio Gonzalez-Martin[2] , Angel A. Juan[2] and Daniel Riera[2]

[1] *Communications Systems (Defence and Space), Indra, Barcelona, Spain*

E-mail: dfernandezp@indra.es

Postal Address: Roc Boronat, 133, Barcelona, Spain.

[2] *Computer Science Dept., IN3-Open University of Catalonia, Barcelona, Spain*

E-mail: {sgonzalezmarti, ajuanp, drierat}@uoc.edu

Postal Address: Roc Boronat, 117, Edifici Media TIC, Barcelona, Spain

Abstract

This paper proposes a heuristic algorithm for solving the scheduling of capacity requests and the periodic assignment of radio resources in a geostationary (GEO) satellite network with a star topology. The network uses the Demand Assigned Multiple Access (DAMA) protocol in the link layer, the Multi-Frequency Time Division Multiple Access (MF-TDMA) as well as the Adaptive Coding and Modulation (ACM) in the physical layer. The proposed algorithm allows processing a given traffic profile with packet expiration time as delay constrains and a maximum tolerated packet loss rate. The processing is completed using the minimum possible spectrum bandwidth. When there is not any structure imposed to the MF-TDMA super-frame, the resource-assignment problem becomes a combinatorial optimization problem which can be seen as a two-dimension (2D) oriented strip packing problem with additional constrains. The well-known Best Fit Decreasing Height (BFDH) heuristic for 2D packing is used as a basis for the proposed allocation algorithm, which is able to obtain a candidate solution in the order of a few hundredths of milliseconds with a commodity computing platform. This real-time response is also necessary condition for most real-life applications of this technology.

Keywords: satellite, DAMA, RRM, BFDH, scheduling, QoS.

1. Introduction

Radio resources management deals with the sharing of radio transmission resources among the users of a radio communication network. A transmission resource provides an amount of communication capacity, which can be used to send an amount of information on a given time period with a given reliability. Depending on the medium access design, a transmission resource can be a chunk of spectrum, a time slot on a chunk of spectrum, a Code Division Multiple Access (CDMA) code, or a combination of them (Açar 2001).

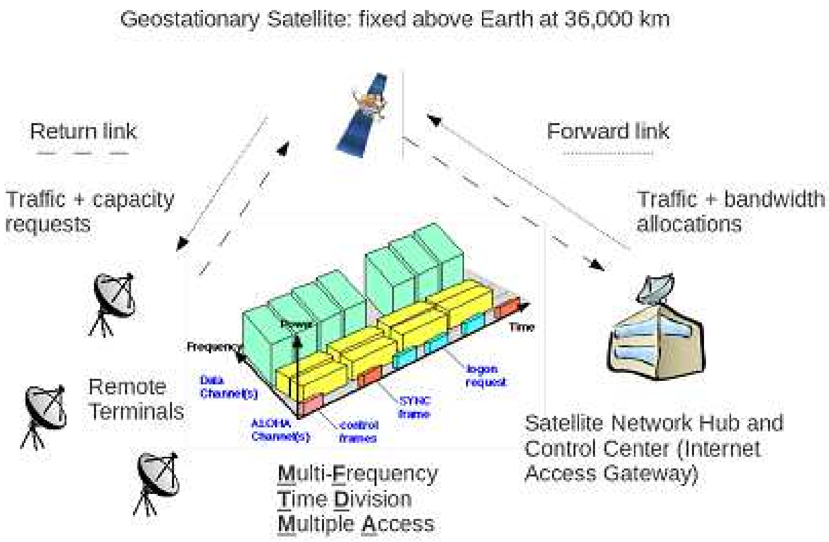

Figure 1: MF-TDMA GEO Satellite Communication System (adapted from Fairhurst 2001)

In a satellite communication network, usually a single hub is the control centre and the responsible of assigning the resources to the multiple user terminals, as shown in Figure 1. In this case, the satellite return link - which goes from the terminals to the hub - is a shared medium that requires a Medium Access Control (MAC) protocol. The Figure 1 shows the satellite communication network considered in the paper, with a star topology and just one hub (the building with an antenna in the rooftop), on the right, assigning transmission resources to the remote terminals (illustrated as three antennas on the left), linked via satellite. In the forward link, from the hub to the terminals, just the hub is transmitting and the remote terminals are all receiving, like TV set top boxes, while in the return link, just the hub is receiving and the remote terminals' transmissions must be coordinated by the hub, so they do not interfere by sending simultaneously on the same time and frequency.

As an example, the proposed MAC protocol by the Digital Video Broadcasting - Return Channel Satellite (DVB-RCS) standard (ETSI 2005) is the Demand Assigned Multiple Access (DAMA) protocol. With DAMA, a Random Access (RA) channel using a protocol such as Slotted Aloha (SA) (Roberts 1975) can be used as a signalling channel to convey small Capacity Requests (CR) towards the hub, where a resource allocation process periodically processes all the received CR from terminals, in order to build a Burst Time Plan (BTP), which indicates the amount of transmission resources assigned to each terminal for its data traffic transmissions until the next BTP reception.

Resource allocation algorithms are not defined within the DVB-RCS standard neither any specified in its guidelines (ETSI 2005). Each commercial DVB-RCS system uses its own confidential algorithms, which are mean to differentiate a DVB-RCS system from the competitor ones.

To contribute solving this lack of open methods, this paper proposes a original heuristic algorithm which allows processing a given traffic profile with messages expiration times (T_D) and a maximum tolerated packet loss rate. The traffic processing is completed using the minimum possible spectrum bandwidth. Since the problem can be seen as a variation of the two-dimension oriented strip packing problem, our algorithm is based on the well-known Best Fit Decreasing Height (BFDH) heuristic. Our approach is able to generate a solution to the periodic resources allocation problem in the order of a few hundreds of milliseconds. This real-time response is also a necessary condition for most real-life applications of this technology.

The structure of this paper is as follows: section 2 introduces some background and concepts on the RRM problem, then performs a problem accurate description and offers a formal model description, in terms of a minimization problem with certain assumptions and constraints. Section 3 performs a literature survey of related papers on the telecommunications field, some of them using packing heuristics. Section 4 explains our approach for solving the problem stated in section 2. It is explained the heuristic general design and it is shown the pseudo-code of the heuristic. Section 5 shows the computational experiments performed with the designed heuristic and the results obtained. Finally, section 6 extracts some conclusions from the work.

2. Background concepts and problem description

Consider a satellite network with one hub and several terminals (Figure 1) using Multi-Frequency Time Division Multiple Access (MF-TDMA). The Radio Resources Management (RRM) problem on the satellite network, introduced by the use of the DAMA protocol, is basically how to periodically distribute available bandwidth and time, which define a Super Frame (SF), for data traffic transmissions among a group of terminals requesting capacity. Depending on the assumptions and constrains considered, the resources usage optimization problem can be formulated in different ways. Hereafter are introduced and clarified some concepts involved with this topic, in order to describe the problem in detail later, in section 2.7.

2.1 Traffic Profile, Quality of Service (QoS) and types of CR

The terminals have to process different types of traffic belonging to different applications. A broadband multimedia network used for Internet access can process a different traffic profile depending on whether the user is a consumer or a professional (prosumer). Moreover, the traffic profile can be different if the satellite network is used for an specific application such as remote monitoring, point-of-sales transactions or aeronautical communications (Air Traffic Management or ATM), instead of Internet and web access.

Voice traffic has been traditionally characterized on telephony networks using a Poisson process, with call durations following an exponential distribution, while data traffic has been characterized using a self-similar model using the Pareto distribution for both, packet sizes and packet inter-arrival times (Becchi 2008). The parameters for the models are usually derived from live traffic captures statistics of packet sizes and inter-arrival times.

Another approach is to use a live traffic packet capture and then create an potentially infinite traffic profile with the same statistical characteristics than the captured sample using the bootstrapping technique (Chernick, 2008), which can be done by numbering the packets in the live traffic packet capture and then using a random uniform distribution for re-sampling the packet capture and indefinitely selecting the next packet that will be generated.

Each type of traffic in the profile can also have some QoS constraints associated. For example, voice must be served with a maximum delay and packet loss to be intelligible. There can be also requirements for the maximum voice call establishment time. While some data applications, such as FTP, email or P2P can be served on a best-effort basis --using the remaining bandwidth after processing traffic with QoS constrains-- interactive applications, such as web browsing, chat or remote consoles have latency constraints (e.g. it can be considered that the great majority of users will give up if a web page takes more than 2 seconds to start loading), so the CR for some data traffic have an associated expiration time, after which it is not useful to allocate the requested capacity, because the connectivity would be considered too bad for the given application and the user would have desisted or switched to an alternative mean to perform the required communication. Some data applications, e.g. network management, can consider also that the maximum expiration time can be achieved sometimes, while a threshold in this number of times is not exceeded.

The terminals of the considered satellite network do not request directly a portion of Super Frame

Bandwidth (SF_Bw) and the transmission time they need. Instead, they periodically report its link condition in terms of signal power and the amount of capacity in bytes needed for each application Class of Service (CoS). In the DVB-RCS standard there are two types of requests defined: volume-based, which correspond to VBDC (Volume Based Dynamic Capacity) CR, and rate-based, which correspond to both, RBDC (Rate Based Dynamic Capacity) and CRA (Constant Rate Assignments) CR. In the DVB-RCS guidelines (ETSI 2005), it is recommended to map rate-based requests to streams of voice and video, while volume-based requests to data traffic. RBDC granted capacity has an associated time-out relatively short, so the resources have to be re-requested periodically, while CRA is granted while the terminal is logged on into the network. Another approach (Mitchell 2004) is to use a high time-out value for rate requests, but then implementing the possibility of sending a capacity release signalling message.

2.2 Spectrum bandwidth

The overall spectrum bandwidth of a satellite network is a fixed parameter, to be manually changed by network configuration procedures. It is determined during the planning of the network, taking into account the foreseen amount of traffic to convey, the capacity provided by the link budget and the modulation and coding techniques used by the satellite modems. It can come imposed also by the available capacity of the satellite used or even by the budget available for spectrum by the satellite network operator. The bandwidth of the satellite network is usually leased to a satellite operator and paid on a monthly basis, used or not. The bandwidth available, jointly with the link budget, determine an upper theoretical limit to the maximum amount of traffic the satellite network will be able to forward, given by the capacity computed by the well known Shannon formula (Shannon 1948).

2.3 Adaptive Coding and Modulation

More and more, with the new satellite systems in Ka-band, Adaptive Coding and Modulation (ACM) is used in order to keep a constant Bit-Error-Rate (BER) in spite of changing link budget conditions, by selecting a more robust MODCOD when Signal-to-Noise Ratio (SNR) is low.

The MODCOD that a terminal uses for transmission is determined by its particular link budget condition on a given time, which depends on the location of the terminal on the satellite coverage zone and also on climate conditions (rain attenuation). The link budget condition can be estimated by the SNR measured by the terminal on the forward link signal from the hub.

Each MODCOD that can be used for transmission provides a different capacity granularity, i.e. an amount of link layer bytes that can be sent on its physical layer burst format. The MODCOD symbol rate determines the amount of bandwidth its bursts use, while the coding and framing structure the link layer bytes that can be conveyed. In order to ease network synchronization, necessary for the TDMA, the rates - so the bandwidths - have values that are multiples of a minimum value. Moreover, it is assumed that in case a burst is not completely filled by user data, padding must be performed.

2.4 Superframe structure

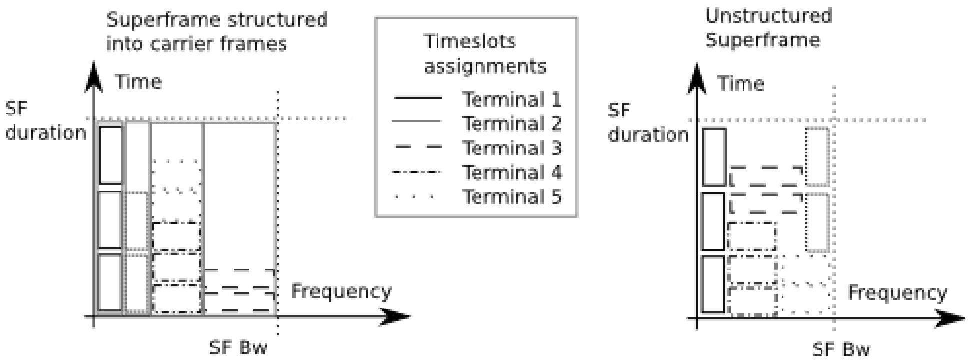

Figure 2: Structured and unstructured superframes timeslots assignments

As shown in Figure 2 left side, the satellite network links bandwidth (Bw) are usually structured and subdivided in carrier frequencies, which can have the same or different bandwidth. Each carrier frequency can contain slots of the same or different durations, due to the use of the same or different

MODCOD (MODulation and CODing) schemes. The sequence of timeslots of a given carrier is called a frame. The group of frames of the different carriers forms the link Super Frame (SF).

Figure 2 right side shows that it could be possible to define an unstructured SF, i.e. not divided into frames, where timeslots of different bandwidths and durations are packed. This scheme leads to allocation algorithms more complex computationally. It can also lead to an increase in the signalling data needed to transmit the BTP, which is better to minimize to keep overall satellite network spectrum utilization low, considering both traffic and signalling in the return and the forward links. For the description of the BTP it is necessary to send: the frequency, bandwidth, starting time and duration of each burst allocated to each terminal. On the other hand the unstructured SF approach allows the maximum optimization regarding the overall bandwidth needed, as shown in Figure 2. Because of this an unstructured SF is considered in this paper, but in order to minimize the signalling required, the proposed allocation algorithm gives preference to resource allocations contiguous, in time, and at the same frequency and bandwidth, trying to join the transmissions from the same terminal as much as possible.

2.5 Superframe duration and assignments period

Although in order to reduce the amount of bandwidth for the RA signaling channel, different allocation periods can be considered according to the different applications' Classes of Service (CoS), with different delay requirements: voice, chat, email, bulk data transfers... usually, because of simplicity, the assignments period (TA) is equal to the SF period (SF_T) for all traffic from all applications.

There is a trade-off in the assignment of a period to DAMA capacity assignments. The larger the period, a more efficient assignment can be performed, using more sophisticated placing algorithms, and with less signalling messages (CR), but the responsiveness of the satellite network will be lower, which can make a bad user experience for interactive applications, such as database requests, web browsing, chat or interactive remote consoles.

On the other hand, in case that the SF_T is small compared to the duration of the frames to allocate, it may make no sense to allocate resources using a complex heuristic and the allocation will be less efficient.

The possible options for the request-assignment cycle period are:

- No period, the assignments are done on-line and published as soon as CR are received. This provides the most responsive possible system but is the most inefficient option.

- Fixed period of tenths of ms, e.g. 26 ms or 32 ms, or hundredths of ms, e.g. 100 ms or 200 ms. This period values are typically used on DVB-RCS networks for Internet access, aiming to provide high responsiveness and a minimum efficiency on the assignment of resources.

The period of allocations TA (Time of Allocations) determines the operation period of the resource allocation process, meaning that it is the minimum period in which allocation of resources to the terminals can be modified. In that respect, in case that the resource allocation must ensure specific delay requirements it is convenient that the allocations' period is much below these requirements, so that the resource allocator will have enough flexibility to modify assignments in front of incoming traffic. This is the case considered in this paper, where the traffic profile considered has strict Quality of Service (QoS) requirements in terms of delay. Taking into account a delay requirement given by TD (Time Delay) and that the procedure of requesting resources and the subsequent transmission will involve a minimum propagation delay of 1.5 RTT (Round Trip Time), a first condition to be held is:

$$TA \ll TD - 1.5\ RTT$$

Considering an engineering criterion $TA \leq 0.1(TD - 1.5\ RTT)$

If TD = 4.7 s and RTT = 540 ms, then TA <= 390 ms.

2.6 Other restrictions and characteristics of the RRM problem

Another restriction that must be considered by the resource allocation algorithm is the fact that a terminal has just one transmitter. It cannot send more than one burst simultaneously, i.e. simultaneous transmission on different frequencies is not possible.

The resource allocation process is periodically executed in two phases (Lee 2003) or three if ACM is considered (Lee 2004). First, the requested capacity is computed overall per terminal and CoS and prioritized by some criteria, e.g. giving priority to terminals with worst link conditions, keeping a proportional fairness (considering link conditions or not, i.e. ACM) and giving priority to some CoS (Vazquez 2005) or just by following an Earliest Deadline First (EDF) policy with received CR, in case they include expiration time information (Modiano 1997). In the case of using EDF it can be considered also the volume requested to avoid postponing long messages in favour of shortest ones.

If ACM is used and the SF is structured, the second step is computing the number of carriers of each type needed. This computation can have a longer period than SF_T in case it is a fixed link with a slow fading, which depends only on meteorological conditions (Aroumont 2008) and not on the terminals movement (fast fading).

On the last step, the BTP is built and resources are assigned to terminals and CoS in the process, following at least the mentioned constrains for time and frequency assignments: overall SF structure, bandwidth and duration and not allowing simultaneous transmissions on different frequencies by one terminal.

CRs that cannot be allocated on a given period are left for the next, and eventually they will get more priority then. The case of interest, of course, is when the network is congested and there are always more CR to allocate than the capacity of the SF, but this situation cannot be permanent, i.e. it must happen from time to time, but without entering the case of a constantly growing queue of CR, which would indicate that the network is overloaded, due to oversubscribing or under-dimensioning.

In general, two questions must be addressed when designing a dynamic bandwidth allocation procedure (Morell 2008): first, how much structure is imposed to the multiple access scheme, then, within the given structure, how are resources optimally distributed. As it has been commented in section 2.4, a highly structured approach gives lower degrees of freedom and simplifies the optimization problem. Giving no predefined structure to the MF-TDMA SF, which is the optimum in terms of minimum bandwidth utilization (see Figure 2), leads to an NP-hard combinatorial optimization problem.

2.7. Considered problem accurate description

Consider a satellite network with one hub and several terminals (Figure 1) using an unstructured MF-TDMA SF i.e. without any pre-fixed set of carrier frequencies configured, as shown in the right side of Figure 2. The terminals' modems can use three types of burst for its data and voice messages transmissions (Figure 3), each one with the same area, depending on its link conditions: bad, average and good.

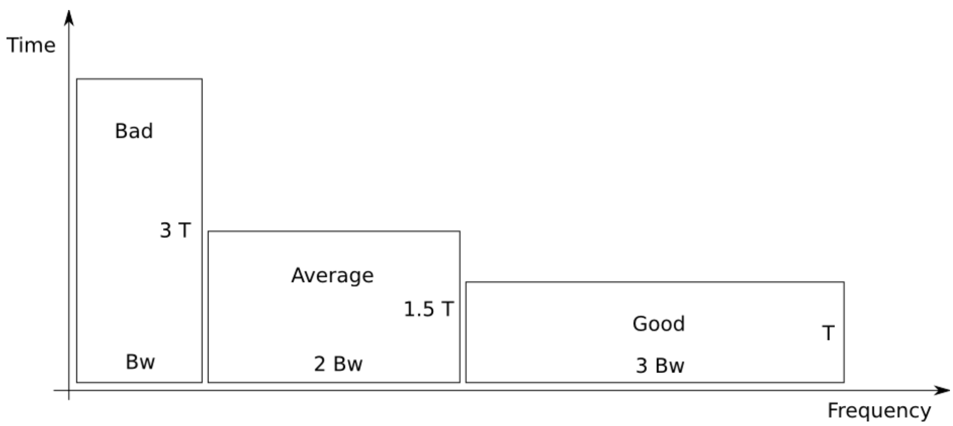

Figure 3: Three types of Radio Frequency (RF) bursts for three channel conditions

A synthetic traffic profile of packets to be processed by the satellite network, using a predefined number of RF bursts with a given timeout has been generated and is available at Fernández 2012. The bootstrapping technique (Chernick, 2008) can be used to generate a traffic profile as long as needed from these generated traces. Traffic profile QoS requirements in terms of latency are specified as expiration times for a group of Radio Frequency (RF) bursts corresponding to a message to be transmitted. It is assumed that all the bursts belonging to a message to transmit must be sent before the overall message expiration time.

The traffic profile determines the types and amount of CR the DAMA allocator process in the satellite network hub will receive by the RA channel used for signalling (CR sending). The dimensioning of the RA channel for CR has been left out of scope of this study, but depending on the characteristics of the traffic profile and the CR strategy implemented on the terminals, it could be a more significant part to the overall bandwidth required in the return link than the DAMA part, which is the objective of this study. In this paper, it is assumed that the input traffic profile represents already received CR at the DAMA allocator process in the hub, with corresponding time-out values updated after its transmission by the CR RA channel, assuming that it is dimensioned so that using more than one retransmission to convey a CR to the hub has a negligible probability.

Figure 4 shows the overall number of bursts to pack of all the three types that appear each SF_T of 390 ms in the most loaded traffic profile considered (file #3).

Figure 5 shows that the CR scheduling problem can be illustrated like a kind of Tetris ™ game. The main input of this problem is the traffic profile, which time axis indicates the time at which a CR for a message transmission need arrives at the allocator process in the satellite network hub. A CR is located at a distance “time-out” of current time and is comprised of a set of bursts to transmit of a given bandwidth and duration, according to one of the three types shown in Figure 3. In the traffic profile it has been assumed that all bursts belonging to a single message are of the same type, i.e. that transmitter MODCOD finally used is known (obtained by a real traffic capture), so there is no need to simulate the ACM process and loop that would be running on a real network and which should be coordinated with the allocator process.

As they are received, the allocator stores CR in order of increasing time-out on a buffer per requesting terminal, where time-out is equal to CR expiration time (TD) minus duration. Notice that when a new CR arrives, due to the constrain that each terminal has just one transmitter, the allocator could have to modify the time-outs of CR already present in just arrived CR corresponding terminal buffer of CR, so there is no time overlap of burst transmissions from a single terminal. It could happen that as a result of the insertion of the just arrived CR some already present CR gets a negative time-out. In this case, it means that the transmission capacity in the short time has been overloaded, resulting in packet loss. This can happen if the scheduling policy is too conservative and the allocator waits too much, almost to the transmission CR time-out expiration, to perform the scheduling of the CR and its corresponding allocation in the BTP, the problem output. One BTP is generated each assignment period that assigns the CR using the minimum possible bandwidth and satisfying constrain of having a single transmitter per terminal.

As assignment periods are done, the expiration time (T_D) of CR pending on the allocator buffers per transmitter approaches. When the deadline is lower than the SF period (SF_T), a CR must be assigned on the current BTP. Otherwise, it would expire the next period and would be lost. The threshold at which a CR deadline must be considered for assignment on a BTP is a configurable input parameter of the algorithm so called obliged transmission threshold. The minimum value of this parameter is the SF period, equal to the allocation period (TA), but there is a trade-off for the value of this threshold, between imposing a limit to the packet losses due to transmitter overload in the considered traffic profile, and the amount of bandwidth required to satisfy the QoS constrains, in terms of delay and packet loss, for the given traffic profile. Despite this trade-off, for a given traffic profile, setting a lower obliged transmission threshold does not necessarily lead to a lower bandwidth demand.

Regarding packet losses due to transmitter overload condition, three overall limits have been considered to obtain results, selected by the value of the obliged transmission (Tx) threshold, shown in Table 1, in terms of PLR (Packet Loss Rate).

CoS	**Expected PLR due to overload**	**Obliged Tx Threshold**
High	0.0 (no losses allowed)	High
Medium	Some losses	(High+SF_T)/2
Low	Potentially maximum losses	SF_T

Table 1 – CoS considered PLR

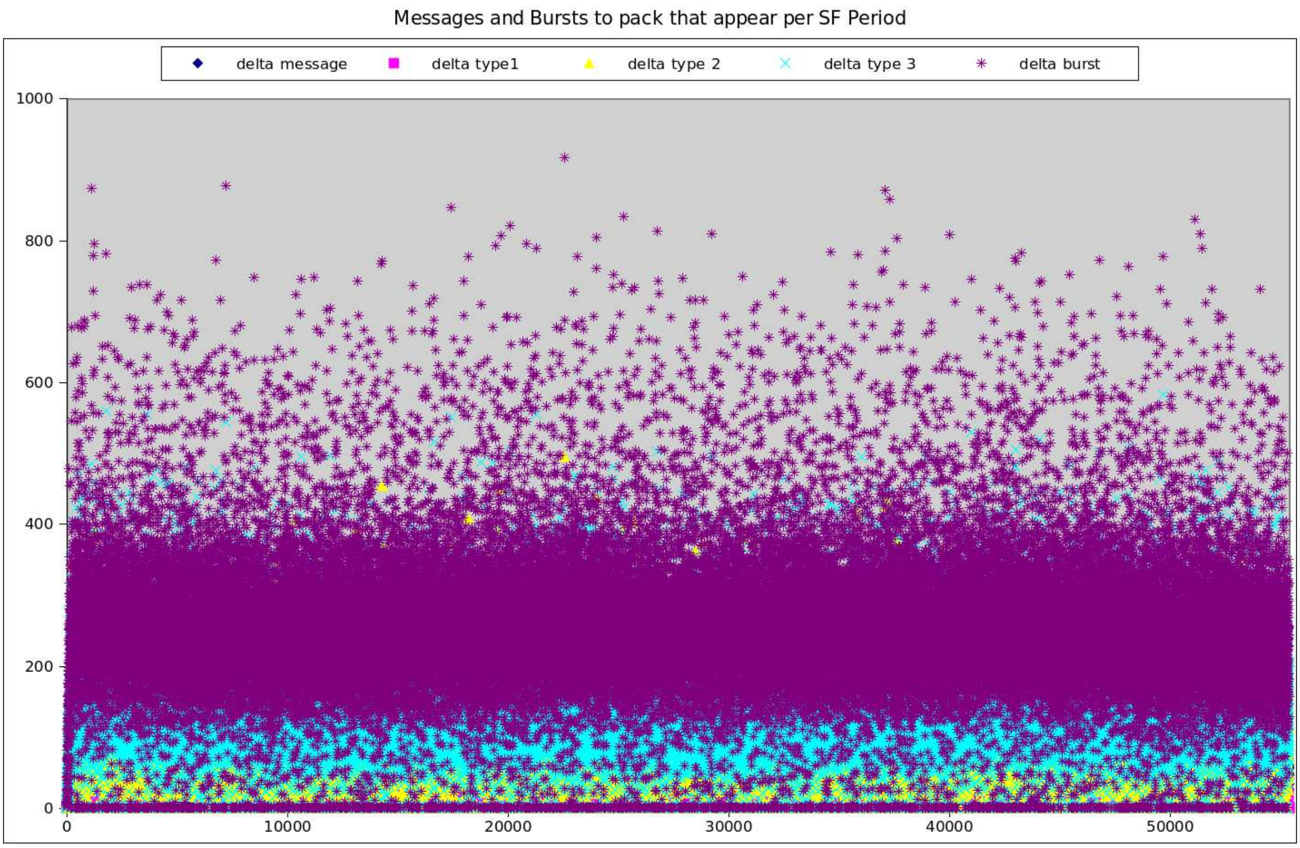

Figure 4: Number of new bursts to pack that appear each assignment period in file #3 (Fernández 2012)

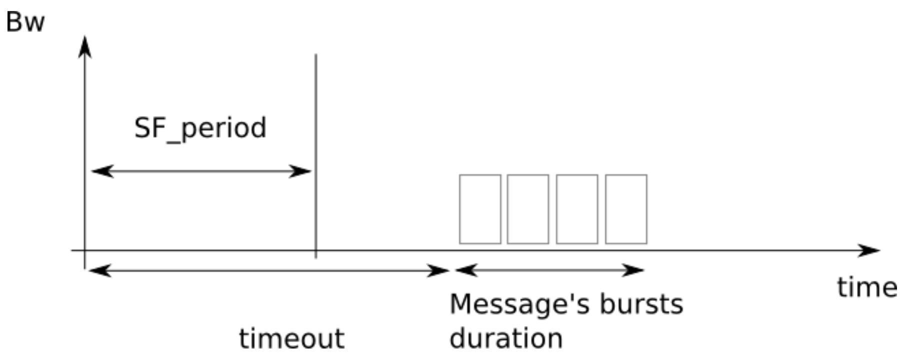

Figure 5: Traffic profile trace parameters illustrated

The RRM problem to solve is to minimize the bandwidth (SF_Bw_min) required to process a given traffic profile set of CR (TPB), with duration T_end seconds, on the described type of satellite communications network. In order to find this minimum bandwidth an allocation algorithm, in this case a bursts packing algorithm (explained later in section 4), has been developed, that solves the next minimization problem:

Minimize SF_Bw(TPB, SF_T, lambda) = SF_Bw_min, subject to:

- Each SF_T seconds, a given SF j has to be filled with a set TPB_j of N bursts b with a given frequency f and time t dimensions from a possible set, see Figure 3, and a time-out T_out each one that is updated each SF_T.

$$TPB_j = \{b_i(f_i, t_i, T\_out_{(i,j)}),\ i=1..N\}$$

$$(f_i, t_i) \in \{(Bw, 3T), (2Bw, 2T), (3Bw, T)\} \Rightarrow SF\_Bw\_min \bmod Bw = 0$$

- T_end = m·SF_T, where m is the number of SF needed to process the traffic profile. T_end is determined by the traffic profile latency QoS requirements.
- SF_T computed from the most stringent data CoS expiration time (TD) and minimum RTT, which constraints the maximum execution time of the allocation algorithm (<< SF_T) and is a multiple of T (shortest timeslot duration), so instead of 390 ms, computed in section 2.5, 360 ms will be considered, assuming T=20 ms:

$$SF\_T \le 0.1(TD - 1.5\ RTT)$$

$$SF\_T \bmod T = 0$$

- All bursts in the profile must be processed. Burst lost (lambda) allowed due to transmissions exceeding the expiration time due to overload, as shown in Table 1.

- A single terminal cannot transmit in more than one frequency simultaneously, i.e. a terminal should not be assigned more capacity than it is able to use.
- In order to minimize BTP signalling size (amount of data required to notify the assignments to the terminals using the forward link) transmitters frequency hopping must be minimized, i.e. it is preferred to send a single transmitter bursts of the same bandwidth in a time row whenever possible.
- The remaining capacity –after performing the allocation of CR exceeding the obliged transmission threshold configured-- is assigned to pending transmissions. This is done to minimize overall latency introduced to the traffic to serve.

In section 4, the approach proposed for solving this problem is explained. The next section 3 shows a literature review of this topic.

3. Related work

3.1 Telecommunications literature survey on RRM

The earliest approaches found to the problem of the optimization resources usage on satellite networks using Bandwidth-on-Demand (BoD) do not consider ACM. They aim at optimizing the resource sharing in case of congestion --i.e. more capacity is requested than resources available-- as an effort to provide the maximum fairness among streams competing for the available resources. The optimization problem is modelled as an integer optimization problem that is derived from Game Theory Açar (2001). Essentially, the solution aims at maximizing the amount of resources allocated to each stream without exceeding the size of the resource pool that is being shared.

In contrast, Priscoli (2004) proposes the decoupling of the congestion control and the BoD mechanism by using control theory concepts (Smith predictor) and modelling the system as a time-delay system.

The terminals make CR in order to track a given maximum reference size of its transmission buffers and avoid over requesting. In case of congestion the max-min fairness criterion is applied, i.e. all users with unsatisfied CR are allocated the same amount of capacity, optionally applying some weights. With this model, the introduction of ACM does not change the algorithm. Bandwidth variations due to changing channel conditions are considered perturbations, in the same way that the ones due to traffic on the network, to be compensated by the control-theoretic algorithm, which tries to keep a reference queue length (Pietrabissa 2008). This reference queue size determines the network utilization and is adapted in function of estimated network congestion, which is performed by monitoring the queue size itself.

Other authors, instead of maximizing the fairness between users, try to maximize the overall throughput of the return link. They model the problem as a linear integer programming problem and try to minimize the sum of the overall differences between requested and assigned capacity, subject to several constrains (Lee 2003). When introducing ACM, the SF is divided into at least two different types of carriers. The problem of maximizing the overall throughput is then modelled as a non-linear integer programming problem (Lee 2004). Then it is not possible any more to use an algorithm to find the exact solution to the optimization problem in the available time, because it becomes NP-complete, so an heuristic is proposed based on the similarity of this problem to the Knapsack problem.

Aroumont et al. (Aroumont 2008) presents an algorithm, which considers again the fairness among users as the optimization criteria, but considering also ACM. The proposed allocation algorithm is based on a water-filling approach. Morell et al. (Morell 2008) consider that the introduction of ACM leads to the need of a cross-layer approach. The resulting RRM problem is cast into a Network Utility Maximization (NUM) problem, which in turn maximizes fairness among users. In that paper it is considered that if no structure is imposed to the SF, the problem becomes NP-hard. Using results from Game Theory, it is proposed that the utility function to maximize when sharing resources among different users is the product of assigned resources.

Note that in these mentioned works the capacity demand constrains are presented from a general perspective --fairness, throughput, etc.-- without relating them to specific requirements in terms of delay bounds --CR expiration times-- and packet losses, as it is the case in this study and also in the ANTARES study (ESA 2011) for Air Traffic Management (ATM) communications. In other approaches, the proposal is not to optimize the use of resources, but just implementing an algorithm, subject to some characteristics, policies and constraints, and then evaluating its performance. This is the case of Mitchell et al. (Mitchell 2004), where the resources are requested on a burst-by-burst basis and assigned following a round-robin policy. This is also the case of Vazquez-Castro et al. (Vazquez 2005), in which paper, terminals with the worst link conditions receive highest priority of allocation.

In Booton (Booton 2008), the DAMA protocol is not performing allocations on an SF basis, as in the previously mentioned works. It is used to allocate or free satellite bandwidth time slots on a more long-term basis. The allocations are done on-line, i.e. as CR arrive, not periodically. In that case, it is also not seen as feasible finding the exact solutions to the optimum allocation. The allocation problem is addressed as a packing problem and Best Fit heuristics are proposed for the problem solution and compared with the previously used First Fit heuristics.

Although not directly related to satellite networking, it is worth mentioning that bin packing algorithm based heuristics are proposed by other authors for the problem of allocating user connections to a set of base stations. In Xing and Venkatasubramanian (Xing 2005) the authors use the First Fit Decreasing heuristic with additional mechanisms in order to accommodate flows, requiring a certain bandwidth and delay QoS constraints, to five available access networks. The bandwidth of each access network is analogous to the capacity of a bin. The bandwidth requirement of a traffic flow is analogous to the size of an item to pack. The objective is reducing power consumption of user terminals while respecting the application preferences, a priori defined by users.

Mariz et al. (Mariz 2006) also draw a parallel with the bounded space variable-size on-line bin packing problem in order to deal with the problem of allocating user services onto a set of communication resources on cooperating access networks. Applications are the objects to be packed and access networks are the bins. There exist a finite number of bins (bounded space) and each of them can have a different size (variable-size). Applications to pack are not known in advance (on-line problem) and the size of applications also depends on the access network they will be finally assigned to. The authors evaluate three well-known bin-packing algorithms: First Fit, Best Fit and Worst Fit.

3.2 2D oriented orthogonal strip packing algorithms

According to (Booton 2008), the BFDH would be the the best heuristic for 2D oriented orthogonal strip packing, at least better than the First Fit heuristic they previously were using.

The Bottom-Left (Chazelle 1983) is another heuristic typically considered for 2D packing problems due to its simplicity. (Lesh 2004) introduces an algorithm that makes a branch-and-bound exhaustive search based on the Bottom-Left heuristic for the 2D oriented orthogonal strip packing problem. Additionally, the algorithm is able to quickly determine if a given set of rectangles can be perfectly packed before running more expensive or less accurate algorithms. It is reported finding the optimum packing for 17 rectangles in less than a second, for 25 rectangles in 2 minutes and the larger problems with 30 rectangles in several hours, on a Linux machine with a 2 GHz processor running non-optimized Java code. Anyway, these times are orders of magnitude higher to the hundredths of milliseconds that we are pursuing (<360 ms).

It mentions also that the most natural permutation to choose for the Bottom-Left heuristic, and the one that works well also in practice, is to order the rectangles by decreasing height (bandwidth in our case), although it would be natural also to try sorting by decreasing width (duration in our case), area (the product of bandwidth and duration in our case), perimeter (in our case a sum of time and duration scalars) and then take the best of the four solutions. It is noticed that, when using the Bottom-Left heuristic, by ordering the set of bursts to allocate by decreasing width (duration in our case) it is guaranteed an allocation with a total height (bandwidth occupation) at most three time the optimum, but the heuristic is not competitive as an approximation when sorted by decreasing height (bandwidth in our case). It is also mentioned that it has been demonstrated that there are examples for which the Bottom-Left heuristic cannot produce the optimum packing under any ordering.

Then, it comments that other heuristics could be considered because of its potential for better solutions, but they would take substantially more time than the Bottom-Left heuristic, such as genetic algorithms or simulated annealing, which are meta-heuristics.

Finally, it explains also their use of the Smallest-Gap heuristic, based on trying to fill the smallest horizontal gap first, which is suspected that slightly, but outperforms the Bottom-Left heuristic. In parallel, a variation of the Bottom-Left, called Left-Bottom, is also tested, and a variation trying to fill the smallest vertical gap first, that are able to solve some cases where the Bottom-Left and the Smallest-Gap heuristics fail.

According to Ntene (2007), off-line 2D oriented orthogonal strip packing problems may be solved using exact algorithms, level heuristics or plane heuristics. In level heuristics the strip is partitioned into horizontal levels according to the tallest item packed on the level. In plane algorithms the strip is not partitioned and items may be packed anywhere in the strip.

In her thesis, 542 benchmark data sets among six categories (Vuuren 2006) are used to compare twenty six algorithms from existing literature: by the average packing height achieved, the frequency with which the smallest packing height is achieved and the execution time. There is a trade-off of execution time with the average packing height. A computerized system was developed using Visual Basic 6.0 that implements all mentioned algorithms in her thesis, which could recommend industry managers the most adequate algorithm for a given packing case. The final recommendation is that if the user is interested in obtaining results rapidly, then the algorithms in the Best Fit and First Fit classes are the best choice.

Imahori (2007) shows an efficient implementation of the Burke algorithm (Burke et al 2004) that requires linear space and O(n log n) time, where n is the number of items to pack. It calls it Best Fit, but it is an heuristic slightly different to the one reviewed by Ntene (2007) as BFDH. This other implementation of Best Fit dynamically selects the next rectangle to place according to available horizontal space and offers two alignment options (tallest or shortest), while in the original Best Fit implementation from 1990 (Coffman 1990), the rectangles are placed strictly according to its decreasing height order and aligned to the left. In general, the Best-Fit heuristic can place thousands of rectangles within 100 ms, but it cannot guarantee a constant approximation ratio in the worst case, as the Bottom-left or the Next Fit, although in practice it is shown that its solutions are just 10% over the optimal. It is worth mentioning that in our case, the rectangles to pack are grouped and prioritized according to its timeout, so the heuristic is completely determining the next item to pack. Moreover, items cannot be rotated, because the time and frequency dimensions are not exchangeable and are determined by the channel conditions, so it is considered that the original Best Fit heuristic implementation (Coffman 1990) matches better our case than the Burke algorithm efficient implementation of Best Fit by Imahori (2007).

4. Our approach

According to our problem description, even though the CRs arrive on-line, the assignments are periodic with the CRs arrived up to the current time, so the bandwidth minimization problem can be modelled as an off-line 2D oriented orthogonal strip packing problem (Ntene 2007), with the mentioned additional restrictions, to solve each assignment period SF_T. The width of the strip is fixed (SF_T) and we want to minimize the height (SF_Bw), by packing at least all the bursts which time-out is less than the configured obliged transmission threshold. This parameter is selected in order to limit the packet loss ratio (lambda) due to transmitter overload events for a given traffic profile. Then, there is the packing of extra bursts that are not going to expiration the next SF_T. The packing of these extra optional transmissions does not lead to the usage of more bandwidth than the currently found SF_Bw_min value. They are added to increase the packing efficiency, minimizing also the overall system latency offered to incoming traffic. The bursts to transmit must be packed up orthogonally and without possibility of reorientation (no rotation allowed), because time and frequency dimensions of bursts are not exchangeable. The packing heuristic levels are piled up in the frequency dimension, without overlapping.

As packing problems are generally NP-hard, it means that it is unlikely that a time-efficient algorithm will be found which is capable to find the optimum solution. This observation directs towards solving the problem using approximate methods such as heuristics. In order to solve this minimization problem, the Best Fit Decreasing Height (BFDH) heuristic for 2D oriented strip packing (Coffman 1990) has been selected as a basis. The BFDH heuristic is an off-line algorithm of greedy type. It obtains the packing by placing the items, from left to right, in rows, forming levels. The first level is the bottom of the bin and upper levels are produced by the horizontal line coinciding with the top of the tallest item packed on the level below. On a given level, an item is packed left justified on that level, among those where it fits, for which the unused horizontal space is a minimum. If no level can accommodate the item, a new level is initialized. BFDH heuristic has been applied to solve different packing problems. Examples are Loni et al. (Loni 2001), which applied it to the one dimensional bin-packing problem (1BP) and the two dimensional bin-packing problem (2BP) and their counterpart one dimensional strip packing (1SP) and the two dimensional strip packing (2SP).

Some changes have been introduced to cope with the additional restrictions and also to simplify the implementation as much as possible, for example, instead of selecting the item which minimizes the remaining area on the level, the item with the biggest area is selected. The proposed heuristic can be executed in a few milliseconds (<<SF_T) in a standard computer. Figure 6 describes the approach using a flow chart, which is explained hereafter.

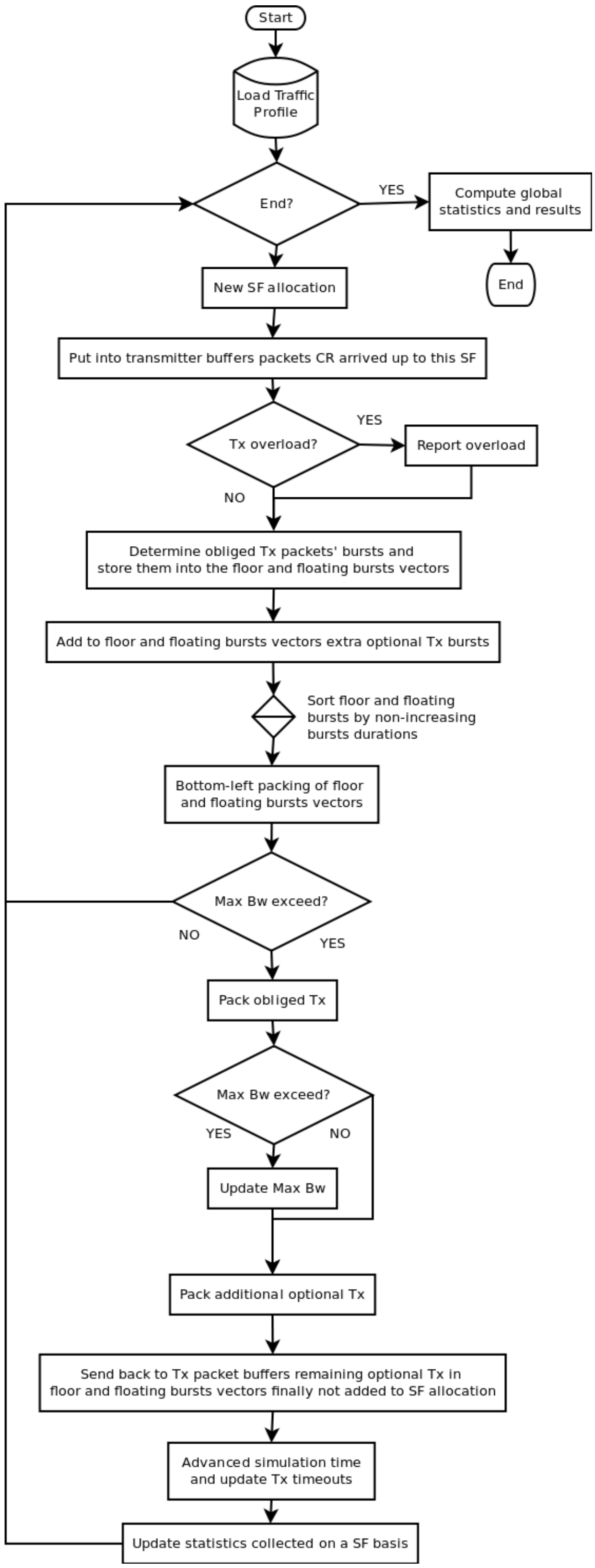

Figure 6: Flow chart describing the overall approach

Our proposed packing algorithm starts by packing first just the obliged transmissions contained in the floor and floating bursts vectors. The floor bursts vector contains the first transmissions of terminals on a given BTP, while if any, the floating bursts vector contains subsequent transmissions of terminals with bandwidth different than the same terminal floor bursts vector transmissions. Each assignment period (SF_T) there is a first stage of collection and aggregation of candidate transmissions from pending CR buffers, with the objective of later minimizing the bandwidth required, but also the signalling required for the transmission of the capacity allocations towards terminals, by minimizing terminals frequency hoping. This first stage is divided in two: the collection of the obliged transmissions first; and then the optional transmissions. The obliged transmission bursts are the ones between the current time and the obliged transmission threshold in the transmitters' buffers, shown in solid colour (not just outlined) in Figure 7. If they are not allocated this SF, the PLR due to overload could be exceeded.

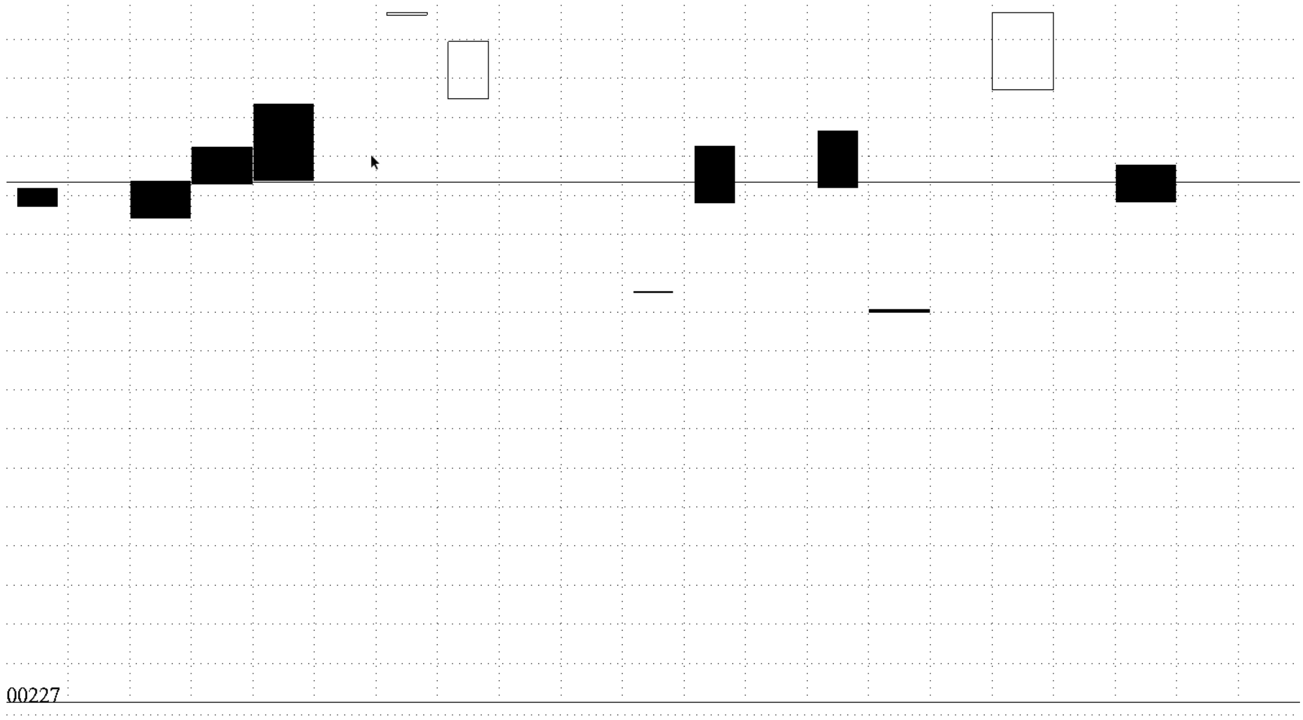

Figure 7: On a given SF period, CRs exceeding the obliged transmission threshold are selected (solid colour), while optional transmissions can be optionally added (outlined)

Then, the optional transmissions are selected following an EDF scheduling policy. The optional transmissions are aggregated to the already existing obliged transmission bursts wherever possible, in order to minimize the signalling required for the allocations.

At the end of this first stage there are filled the two vectors of floor and floating bursts, with both obliged and optional transmission bursts. Note that the fact a burst is floor or floating is determined by the chance of the alignment of the CR arrivals with the period of assignments. On a second stage, the floor and floating bursts vectors are ordered by non-increasing duration of bursts and simply bottom-left packed.

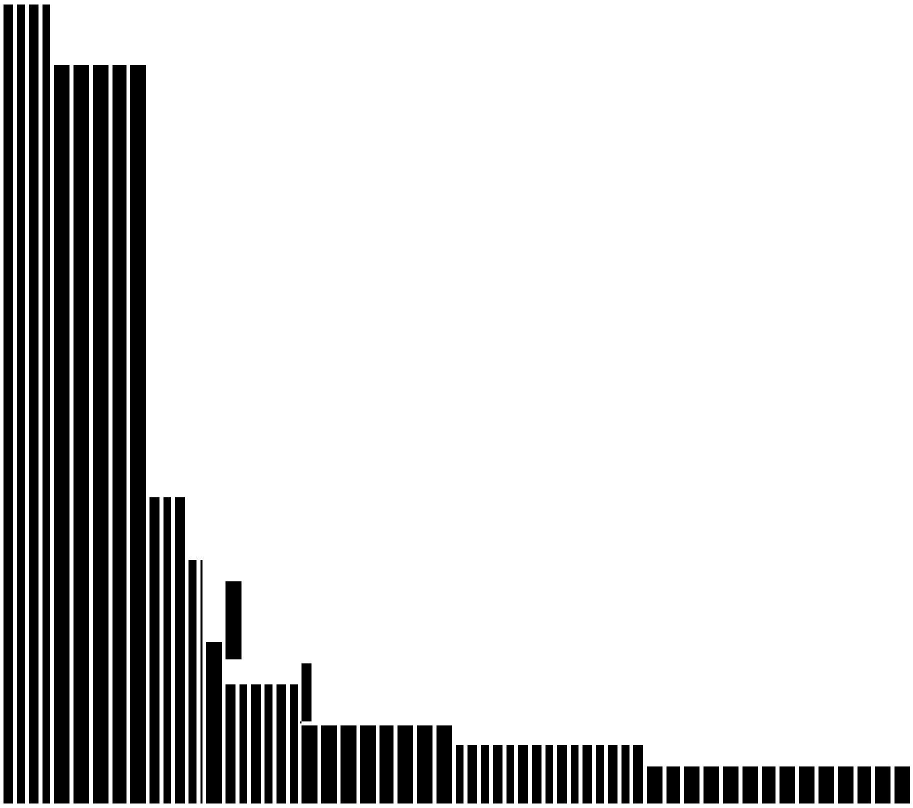

Figure 8: Floor and floating bursts ordered and preliminarily packed

Additionally, obliged and optional candidate transmission bursts are blended in the floor and floating bursts vectors ordered by non-decreasing duration. The devised heuristic adapting BFDH has been compared to this simple approach of bottom-left packing transmissions, in order to check the improvement achieved of the proposed algorithm against a do nothing approach, i.e. serve requests as they arrive just bottom-left packing them. In case this initial simple packing, shown e.g. in Figure 8, exceeds the maximum bandwidth used by all SF up to now or initially configured, the bursts in the floor and floating bursts vectors will be repacked more efficiently using our proposed heuristic, based on the BFDH heuristic.

Similarly to the original BFDH heuristic, levels are created where bursts are packed, but in our proposal, in order to minimize the signalling required, each level must pack bursts of the same bandwidth. Another variation from the classical heuristic, introduced to simplify the implementation, is that instead of selecting for packing the item that leaves less residual empty space on any level, the biggest burst in terms of area that can be packed on an existing level is selected. If there is not found any candidate burst to be packed, the next biggest area burst in the floor bursts vector is selected and packed on a new level. The process continues until all obliged transmissions in the floor and floating bursts vectors are packed into current SF. As a result we get a bandwidth occupation, which can be higher than the maximum used up to now or not. In case it is higher, the value of SF_Bw_min is updated, which at the end of the traffic profile processing computations results in the minimum bandwidth that the satellite communications network operator has to rent to a satellite operator for the return link DAMA, in order to strictly satisfy the QoS needs of the considered input traffic profile.

After obtaining this minimum bandwidth required result, extra packing of bursts can be done in order to occupy the SF as much as possible and to obtain a better system latency and available resources utilization efficiency. The packing of additional optional bursts continues as far as there are candidates found that can be packed into existing levels, but, of course, no new levels can be created to pack optional transmissions, which would lead to more bandwidth required.

Figure 9 shows an example of an allocation of obliged (filled) and optional (outlined) burst transmissions on a SF, and the remaining empty space.

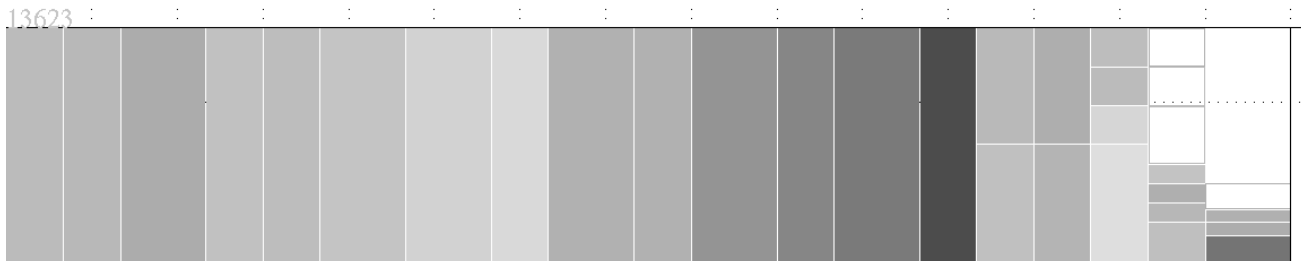

Figure 9: An example of an allocation minimizing bandwidth and signalling required

Bursts selected as optional transmissions but finally not allocated are sent back from the floor and floating bursts vectors to its corresponding packet transmitter buffers, in the same place they were before the allocation process, but with their deadlines updated one SF period.

Then, the appeared message bursts transmission needs (CR) are positioned in its transmitter buffers and the process explained is then repeated for the next to come SF period.

Note that the algorithm has been implemented with the purpose of system dimensioning. It is implemented as a procedure to determine the minimum bandwidth required to process a given traffic profile using as much bandwidth as needed. In order to be used on a real system for resources assignment, it should be slightly modified, to perform the packing the same way, but discarding packets when bandwidth initially configured is not enough, instead of increasing it.

The packing algorithm outlined in Figure 6 steps in case maximum bandwidth is exceeded, are explained here after in more detail using pseudo-code notation. It is based on the BFDH pseudo-code from (Ntene 2007), in which our algorithm is also based.

As it has been mentioned, our proposed packing algorithm needs to perform two steps, one to pack the obliged transmissions, which determine the minimum bandwidth required, and a second one to pack additional optional transmissions that can be advanced to maximize packing efficiency and minimize system latency. Moreover, in order to minimize signalling, interferences and to ease network synchronization in a real system implementation, bursts are packed only in levels with the same bandwidth for already present bursts. Otherwise, a new level is created.

Figure 10 shows the pseudo-code of the *OptimizeBandwidth* method, which implements the BFDH heuristic for a given SF. This method receives as parameter the list of floor bursts, which are the first bursts of any terminal which is willing to transmit during the current SF, and also receives the list of floating bursts which are the subsequent bursts with a different bandwidth, if any. First of all, the method sorts the bursts of both lists by decreasing duration. Next, it creates the solution for the SF allocation and starts by packing the first floor burst. After that, it starts a loop for packing every one of the floor bursts. It is checked at each allocation if there is enough free space in the solution to pack the bursts. If so, the burst is packed, otherwise another level is created on the solution and the burst is packed. Following this loop, another loop is started to execute the same procedure with the floating bursts list. Once it is finished, the obtained solution is returned.

```
procedure OptimizeBandwidth(floorBursts, floatingBursts)
sortByDuration(floorBursts);
sortByDuration(floatingBursts);
solution = new Solution();
solution.packRectangle(floorBursts[0]);
for i=1,…,floorBursts.Length()-1 do
  level = searchLevelsWithEnoughSpace();
  if level then
    solution.packRectangle(floorBursts [i], level);
  else
    newLevel = solution.createLevelAboveTopMostLevel();
    solution.packRectangle(floorBursts [i], newLevel);
  endif
endforeach
for i=1,…,floatingBursts.Length()-1 do
  level = searchLevelsWithEnoughSpace();
  if level then
    solution.packRectangle(floatingBursts [i], level);
  else
    newLevel = solution.createLevelAboveTopMostLevel();
    solution.packRectangle(floatingBursts [i], newLevel);
  endif
endforeach

return solution;
end
```

Figure 10: Pseudo-code of the *OptimizeBandwidth* method

5. Numerical experiments

The resource allocation process described in the previous sections has been implemented as a C/C++ console application, GPLv3 licensed and available from http://sourceforge.net/projects/trafficprofilep. The implementation of the packing heuristic pseudo-code shown in previous section has been inspired on Java code from the Burke Best Fit implementation of the Two- Dimensional Loading Capacitated Vehicle Routing Problem (2L-CVRP) problem by Juan 2012. The experiments have been executed over a standard laptop with an Intel® Core™ 2 Duo CPU at 2.4 GHz and 4 GB of RAM.

To evaluate the algorithm, five different traffic profiles have been generated. In Table 2 are listed the results obtained for each file in the traffic profile if bandwidth is infinite, i.e. a message can be transmitted as soon as it arrives, and just bottom-left packing of transmissions as they arrive is performed.

File #	Max. Bw Used	SF# Max. Bw	Max. Num. Active Tx
1	20	52356	10
2	32	3569	16
3	247	46014	97
4	114	10357	49
5	43	56976	19

Table 2 – Results for each traffic file considering bottom-left packing

The reported figure "Max num of active Tx" refers to the maximum number of terminals with queued packets, not to the maximum number of carriers on a SF. The "Max Bw used" figure units are not Hz, but multiple of Bw, the bandwidth of the narrowest burst (in terms of bandwidth) shown in Figure 3.

Table 2 allows comparing each file in terms of traffic volume contained. There are mainly two types of files: #1, #2 and #5, with a lower load, and #3 and #4 with a higher load.

On the next experiment, our packing heuristic for capacity allocation when no packet losses are admitted was tested. In this case a message could not be transmitted as soon as it was received, but should be transmitted before the obliged transmission (Tx) threshold time elapsed. Table 3 shows the results of this experiment. There are listed the bandwidth needed values obtained with our packing heuristic for capacity allocation when no packet losses are admitted, and the resulting obliged Tx thresholds considered, which seems that mainly depends on the traffic profile of voice calls present in the traffic file.

File #	Max. Bw Used	SF # Max. Bw	Max. Num. Active Tx	Obliged Tx Threshold (ms)

1	9	178	33	777
2	15	21236	60	1144
3	86	42469	1145	3487
4	45	8461	288	12567
5	21	227	39	5168

Table 3 – Results for each traffic file considering limited bandwidth available and PLR=0

Table 4 compares the resulting needed bandwidth in the case of limiting bandwidth used for PLR=0 to the case of having infinite bandwidth.

File #	**Max. Bw Used (PLR=0, finite Bw)**	**Max. Bw Used (infinite Bw)**	**Reduction %**	**Average SF Resources Utilization %**	**Max. SF Resource Utilization %**
1	9	20	55	9,79	100
2	15	32	53,13	16,13	100
3	86	247	65,2	45,1	100
4	45	114	60,53	38,32	96,3
5	21	43	51,16	16,97	96,83

Table 4 – Bandwidth reduction obtained by our algorithm with respect to simple bottom-left packing

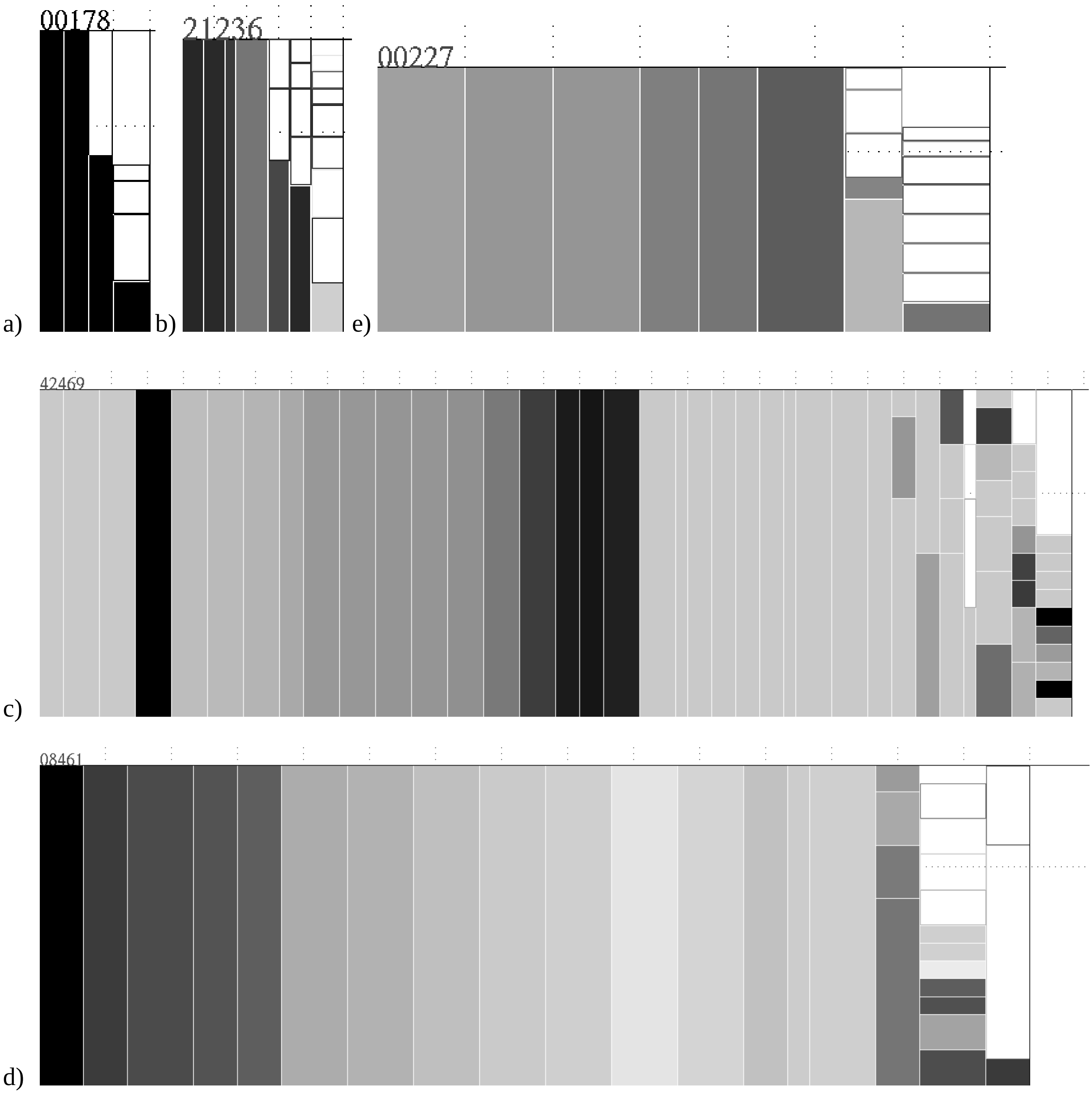

Figure 11: Allocations performed in the SF of maximum bandwidth usage for PLR=0 for files 1(a), 2(b), 3(c), 4(d) and 5(e)

The console application generates some text files that can be used by an accompanying OpenGL based visualization program in order to plot the packing performed each SF period, as shown in Figure 10, which shows the allocations performed in the SF where the maximum bandwidth needed was reached for each input file, as listed above.

In order to discuss the efficiency of our solution, the packing efficiency, measured as area used vs. area available has been obtained on a SF_T basis for each file and packet loss case considered. In the measured efficiency values shown in Figure 12, Figure 13, Figure 14, Figure 15 and Figure 16 hereafter, the minimum resulting bandwidth needed was fixed since the beginning of the software execution.

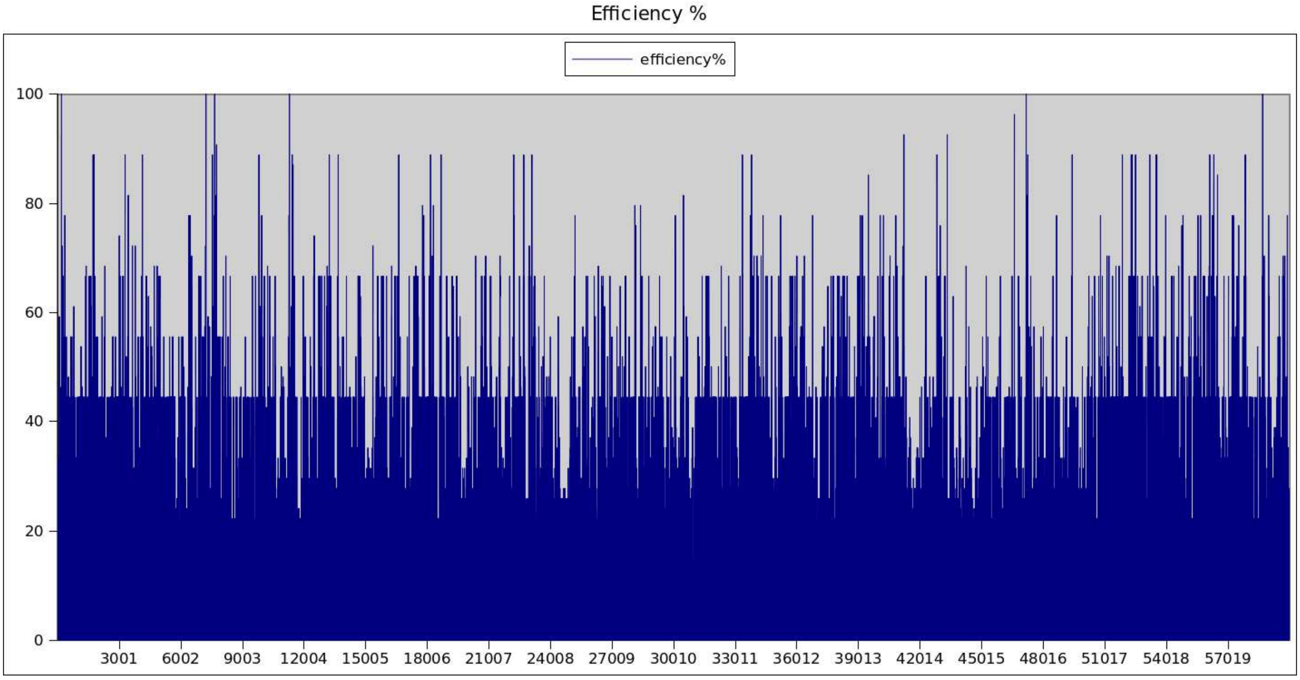

Figure 12: Packing efficiencies measured each SF for file #1 setting SF Bw = 9 since the beginning (target PLR = 0)

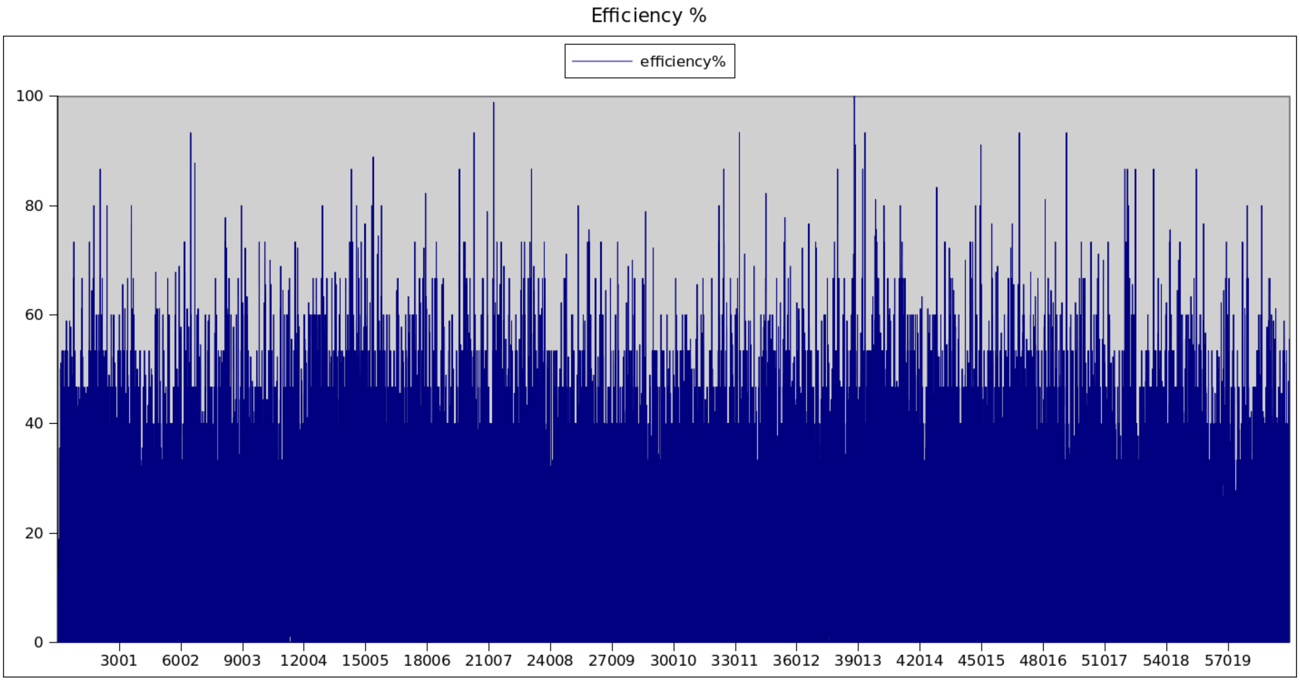

Figure 13: Packing efficiencies measured each SF for file #2 setting SF Bw = 15 since the beginning (target PLR = 0)

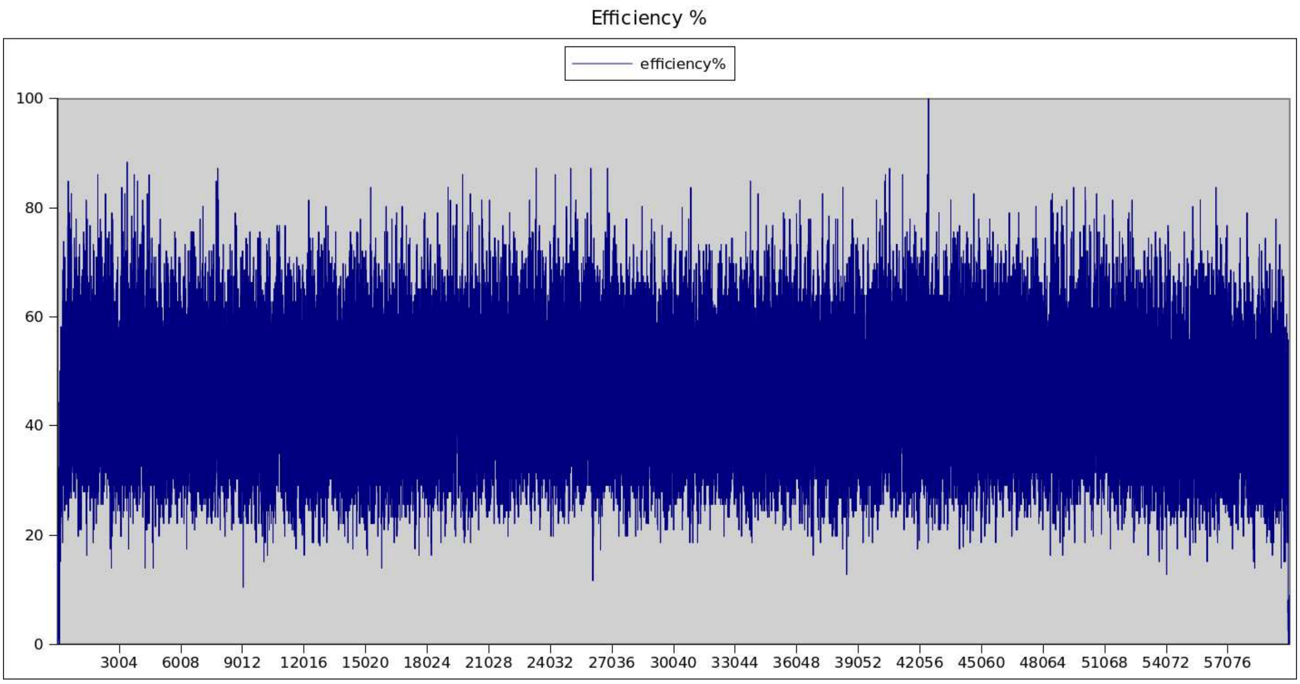

Figure 14: Packing efficiencies measured each SF for file #3 setting SF Bw = 86 since the beginning (target PLR = 0)

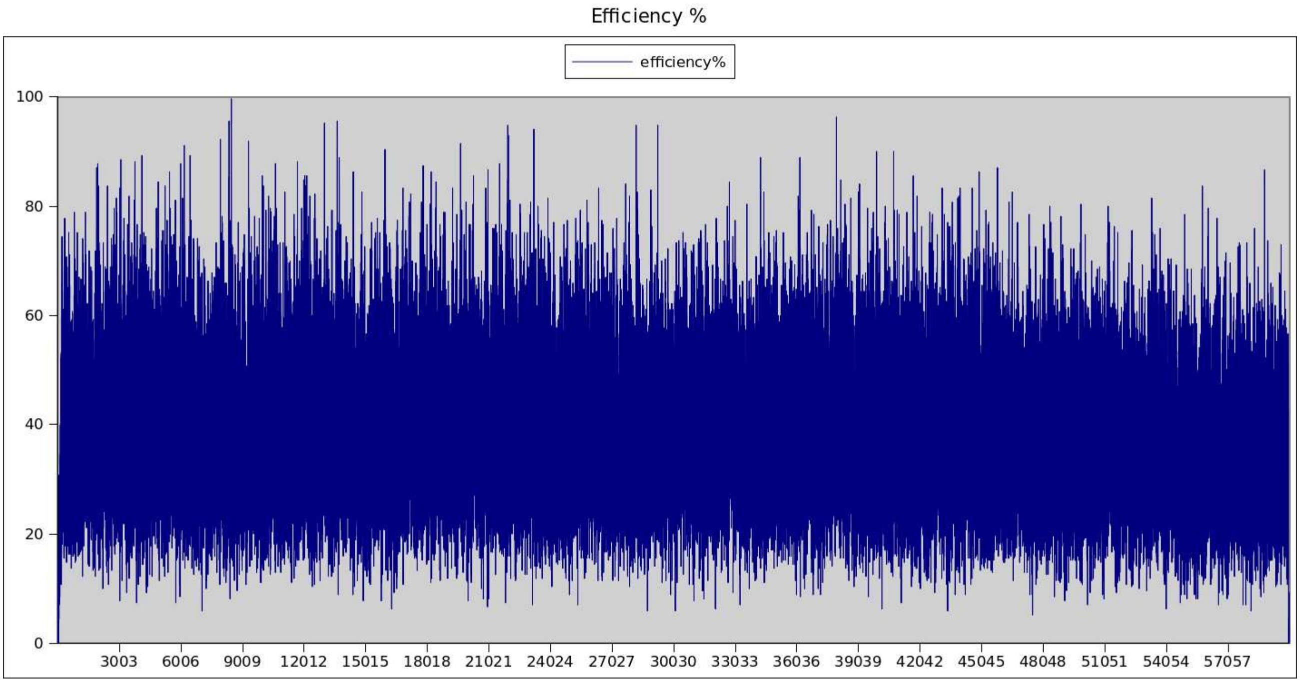

Figure 15: Packing efficiencies measured each SF for file #4 setting SF Bw = 45 since the beginning (target PLR = 0)

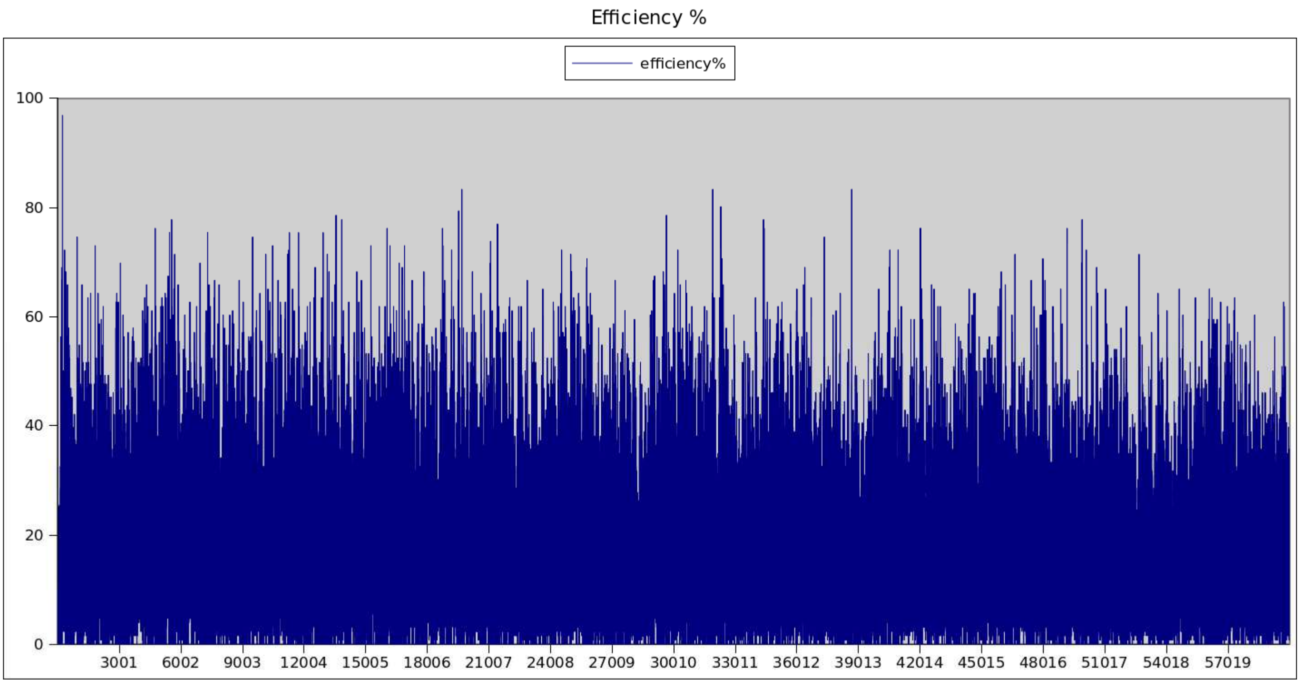

Figure 16: Packing efficiencies measured each SF for file #5 setting SF Bw = 21 since the beginning (target PLR = 0)

As a general conclusion, it is remarkable the fact that > 95% packing efficiency is reached some times in all files. The packing efficiencies and the maximum required bandwidth depend a lot on the considered traffic profile, which is quite in bursts in the five considered instances, so the average-case efficiency measured is relatively low, but it needs to be this way in order to satisfy the applications required QoS at a few critical moments. The more traffic a file contains the greatest the average efficiency is, due to the greatest possibility to exploit an statistical multiplexing of traffic.

Table 5 shows a rough estimation of the amount of signalling needed in the forward direction to convey allocations to terminals. It can be obtained by filtering, e.g. using the serial editor sed (sed ’1~2d’ display_file.txt > assignments_file.txt), the trace files describing the allocations performed each SF period, used by the OpenGL based viewer, that have been used to generate the plots of packing each SF period, see, e.g. Figure 11. The average rate is obtained just dividing the amount of signalling by the six hours that last the traffic traces. The minimum peak rate, can be obtained by diving the longest line size in characters (bytes) - obtained e.g. with wc -L < filename.txt - describing the assignments in the files, by the SF_T. Another option is to dimension the signalling channel capacity for the worst case, i.e. a SF filled with the smallest bursts and all from different terminals, to be notified in just an SF_T.

File #	**SF Bw at t0**	**Amount of Signaling**	**Minimum Peak Signalling Capacity (bits/s)**
1	9	2.0 MB (2069001 bytes)	6200
2	15	4.0 MB (4190110 bytes)	9778
3	86	53.2 MB (55767273 bytes)	34956
4	45	28.5 MB (29927956 bytes)	21534
5	21	6.4 MB (6754431 bytes)	9712

Table 5 - Amount of signalling and minimum signalling channel capacity needed in the forward direction to convey allocations to terminals

Moreover to the higher computational time, one of the main cited drawbacks of following a non-structured approach for SF resources assignment is the high amount of signalling to send on an SF period basis. But it must also be considered that the satellite links are quite asymmetric. The capacity in the forward direction (from hub to satellite terminals) is usually much higher than in the return link direction. Whether or not the amount of signalling required by the proposed heuristic is acceptable or excessive will depend on the characteristics of the considered communication network, but it is an important point to take into account in the design trade-offs of the satellite network return and forward links.

Hereafter are compared the results obtained for each file when the obliged transmissions threshold is set to the other values considered in Table 1, which are the lowest possible value, i.e. SF_T and to a value between the minimum and the previously reported values for PLR=0 (see last column of Table 3).

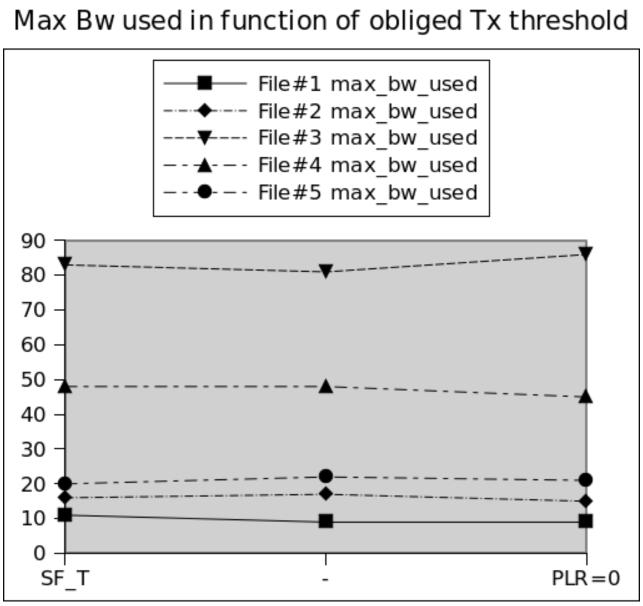

Figure 17: Maximum bandwidth used in function of obliged Tx threshold

Looking at Figure 17, there does not seem to be a dependency between the resulting maximum bandwidth used and the value set for the obliged transmissions threshold. The same can be said of efficiency, according to Figure 18.

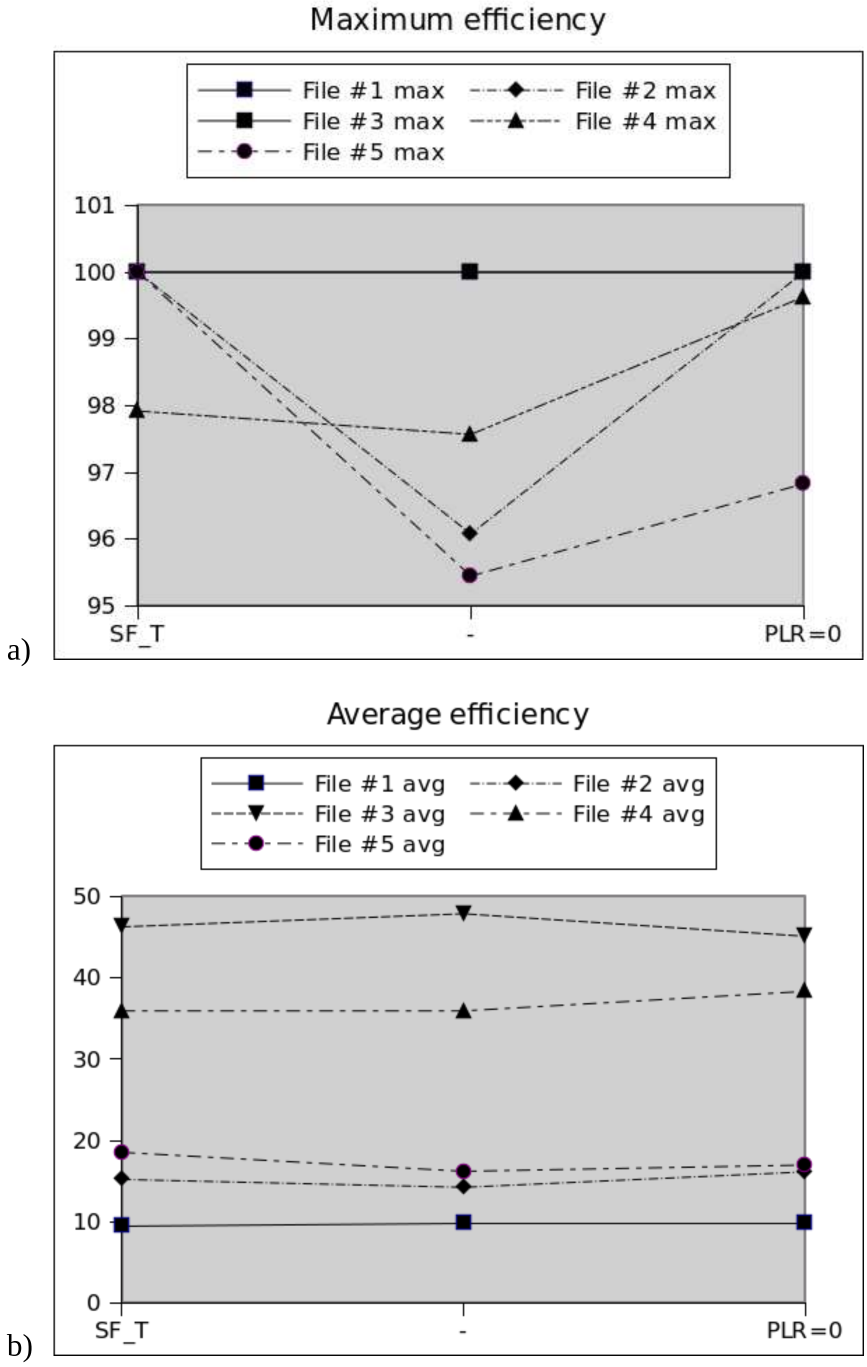

Figure 18: Maximum (a) and average (b) resources utilization efficiency in function of obliged Tx threshold

PLR due to overload

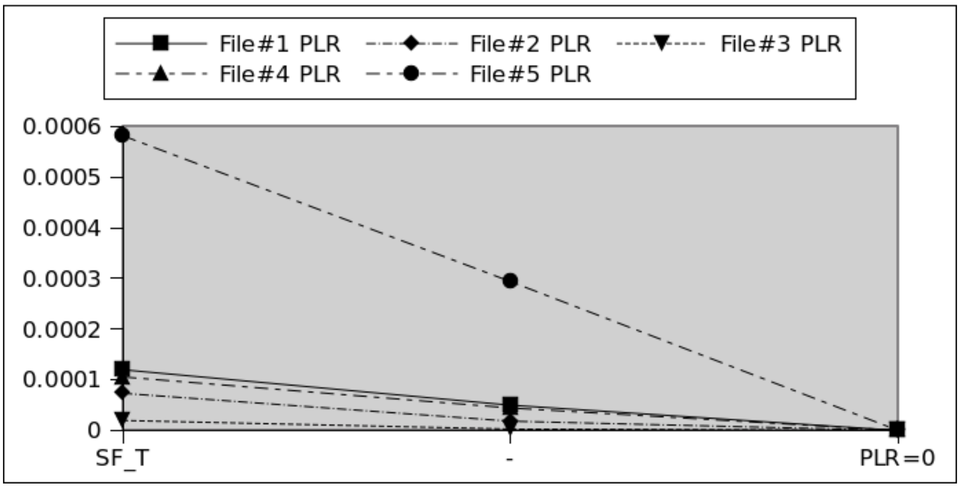

Figure 19: PLR due to transmitter overload in function of obliged Tx threshold

The packet losses due to transmitter overload always decrease as the obliged transmission threshold increases, as shown in Figure 19. It has been found, by tracing the software, that the transmitter overloads are due to the expiration of streams of packets that must be transmitted continuously, without any gap time in the transmission, and which arrive at a rate exceeding the maximum transmission rate, e.g. arrival of 60 ms duration bursts each 20 ms.

6. Conclusions

This paper has introduced the analysis of a resource allocation problem on satellite networks which was never studied with strip packing techniques. Previous approaches for this problem consisted in defining the required bandwidth in a very conservative way or using very simple allocation methods such as bottom-left packing, so that the final system was designed with a bandwidth much greater than what in fact was required.

This paper has also shown the design and performance of a new heuristic to solve the resource-assignment problem on communications networks, assuming an unstructured approach to the problem, i.e. without a predefined set of carriers defined. It is based on the Best Fit Decreasing Height (BFDH) heuristic for 2D packing problems. This is a novel and innovative approach to the Radio Resources Management (RRM) in telecommunication networks in two dimensions. First, because unstructured approaches are frequently discarded in telecommunications literature due to computational cost, in favour of more structured approaches, easier to implement and characterize. But this paper shows that the unstructured approach to the problem is feasible for considered representative traffic profiles, without the need of extraordinary computing resources. On the other hand, it is well-known that innovations usually happen by merging two or more separate topics. This paper algorithm is quite innovative in its approach of solving a telecommunications networks problem from an operations research problem perspective, adapting an existing heuristic from the field of strip packing problems (BFDH) to the telecommunications networks resources management, by establishing an analogy between orders of strips of materials and the need to transmit Radio Frequency (RF) bursts packed on a frame of time and spectrum.

7. Acknowledgments

This work has been partially supported by the Spanish Ministry of Science and Innovation (TRA2010-21644-C03) and by the Ibero-American Programme for Science, Technology and Development (CYTED2010-511RT0419) in the context of the ICSO-HAROSA Programme of the IN3 http://dpcs.uoc.edu).

8. References

Açar, G. (2001). "End-to-end resource management in geostationary satellite networks". Ph.D. dissertation, Dept. Elect. Electron. Eng., Imperial Coll., London, U.K.

Aroumont, A., Radzik, J., Bousquet, M., Castanet, L. (2008). "DVB-RCS Return link Radio Resource Management for Broadband Satellite Systems using Fade Mitigation Techniques at Ka band". IEEE International Workshop on Satellite and Space Communications (IWSSC) 2008

Becchi, M. (2008). “From Poisson Processes to Self-Similarity: a Survey of Network Traffic Models”. http://www.cs.wustl.edu/~Jain/cse567-06/ftp/traffic_models1/index.html. Retrieved on April 2012.

Booton, R. (2008). “Improving Satellite Bandwidth Utilization by Applying Combinatorial Optimization to the Integrated Waveform (DAMA UHF Satcom)”. IEEE 2008.

Burke EK, Kendall G. & Whitwell G., (2004). “A New Placement Heuristic for the Orthogonal Stock-Cutting Problem”. Operations Research, 52(4), pp. 655–671.

Chazelle, B. (1983). “The Bottom-Left Bin-Packing Heuristic: An Efficient Implementation”. IEEE Transactions on Computers, Vol. C-32, No. 8, August 1983.

Coffman Jr. EG & Shor PW. (1990). “Average-case Analysis of Cutting and Packing in Two Dimensions”. European Journal of Operational Research, 44, pp. 134–144.

ESA (2011) “ANTARES Project”. http://telecom.esa.int/telecom/www/object/index.cfm?fobjectid=30228. Retrieved on March 2012.

ETSI (03/2005) TR 101 790 v1.2.1 “Digital Video Broadcasting (DVB); Interaction channel for Satellite Distribution Systems; Guidelines for the use of EN 301 790 ”.

ETSI (09/2005) EN 301 790 v1.4.1. “Digital Video Broadcasting (DVB); Interaction channel for satellite distribution systems”.

Fairhurst, G. (April 2001). “Broadband Satellite Overview”. http://www.abdn.ac.uk/erg/research/future-net/digital-video/bband-sat.html. Retrieved on March 2012.

Fernández, D. (2012). Traffic Profile Processing. http://sourceforge.net/projects/trafficprofilep/ Retrieved on July 2012.

Imahori, S., Mutsunori, Y., (2007). “The Best-fit Heuristic for the Rectangular Strip Packing Problem: An Efficient Implementation”. Mathematical Engineering Technical Reports, September 2007. Dept. of Mathematical Informatics. Graduate School of Information Science and Technology. University of Tokyo. http://www.keisu.t.u-tokyo.ac.jp/research/techrep/data/2007/METR07-53.pdf Retrieved on May 2012.

Juan A. A., Domínguez, O., Faulin, J., Agustin, A. (2012). “Multi-Start Algorithm for Solving the Two-Dimensional Vehicle Routing Problem with Items Rotation ”, International Transactions in Operational Research.

Lee, K-D., Chang, K-N. (2003). “Throughput-Maximizing Timeslot Scheduling for Interactive Satellite Multiclass Services”. IEEE Communications Letters, Vol. 7, No.6, June, 2003, pp. 263-265

Lee, K-D., Chang, K-N. (2004). “A Real-Time Algorithm for Timeslot Assignment in Multirate Return Channels of Interactive Satellite Multimedia Networks”. IEEE Journal on Selected Areas in Communications, Vol. 22, No.3, April, 2004, pp. 518-528

Lesh, N., Marks, J., McMahon, A., Mitzenmacher, M., (2004). “Exhaustive approaches to 2D rectangular perfect packings”. Information Processing Letters, Volume 90, Issue 1, 15 April 2004, Pages 7-14.

M. Chernick. Bootstrap Methods: A Guide for Practitioners and Researchers. Wiley Series in Probability and Statistics. Wiley, 2008 .

M. S. Loni, A. and M. Monaci. Two-dimensional packing problems: A survey. European Journal of Operations Research, 141, pages 242–252, 2001.

Mariz, D., Cananea, I., Sadok, D., Fodor, G. (2006). “Simulative analysis of access selection algorithms for multi-access networks”. WOWMOM'06: Proceedings of the 2006 International Symposium on on World of Wireless, Mobile and Multimedia Networks. Washington, DC, USA: IEEE Computer Society, 2006, pp. 219-227.

Mitchell, P. D., Grace, D., Tozer, T. C. (2004). “Burst Targeted Demand Assignment Multiple-Access for Broadband Internet Service Delivery Over Geostationary Satellite”. IEEE Journal on Selected Areas in Communications, Vol. 22, No.3, April, 2004, pp. 546-558

Modiano, E. (1997). “Scheduling packet transmissions in a multi-hop packet switched network based on message length”. Proceedings, Sixth International Conference on Computer Communications and Networks, Las Vegas, NV , USA.

Morell, A., Seco-Granados, G., Vázquez Castro, M.A. (2008). “Cross-Layer Design of Dynamic Bandwidth Allocation in DVB-RCS”. IEEE Systems Journal, Vol. 2, No.1, March, 2008, pp. 62-73

Ntene, N. (2007). “An Algorithmic Approach to the 2D Oriented Strip Packing Problem”. Ph. D. dissertation, University of Stellenbosch, South Africa.

Pietrabissa, A., Baudoin, C. (2008). “Radio Resource Management for Next Generation DVB-RCS Systems”. IP Networking over Next-Generation Satellite Systems, 2008, Springer, pp. 91 – 101

Priscoli, F., Pietrabissa, A. (2004). ”Design of a Bandwidth-on-Demand (BoD) protocol for satellite networks modelled as time-delay systems”. Automatica 40, pp. 729-741.

Roberts, L. G., Lawrence G. (April 1975). "ALOHA Packet System With and Without Slots and Capture". Computer Communications Review 5 (2): 28–42

Shannon, Claude E. (July/October 1948). “A Mathematical Theory of Communication”. Bell System Technical Journal 27 (3): 379-423.

Vázquez-Castro, M.A., Ruggiano M., Ronga,L.S. and Werner, M. (2005). “Uplink Capacity Limits for DVB-RCS Systems with Dynamic Framing and Adaptive Coding”. Proc. 23rd In. Com. Sat. Sys. Conf., Rome, Sep 2005.

Vuuren (2006). 2D Strip Packing Problem Benchmark Data Sets. http://dip.sun.ac.za/~vuuren/repositories/levelpaper/spp%5B1%5D.htm. Retrieved on March 2012.

Xing, B., Venkatasubramanian, N. (2005). “Multi-constraint dynamic access selection in always best connected networks”. MOBIQUITOUS'05: Proceedings of the The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services. Washington, DC, USA: IEEE Computer Society, 2005, pp. 56-64.