Effective VoD Service Distribution over Cable Television Networks

Lidia Jordanova1† and Jordan Nenkov2††,

Technical University of Sofia, Bulgaria

Summary
The paper deals with the concept for realizing a VoD system over cable television (CATV) networks in a way to ensure possibilities to increase the number of both the subscribers and the movies supported. An algorithm for the effective movie content distribution is suggested that makes the system reliable. Statistical investigations have been carried out to determine the distribution laws of the movies content duration, the file size and the drop in popularity of the series offered. The obtained results are discussed.

Key words:
VoD system, HFC CATV network, Distribution algorithms, MPEG-2 TS, Movie file size

1. Introduction

CATV networks deployment is a serious problem cable operators have to tackle in order to provide high quality services to their subscribers and avoid the necessity of adding new video servers too often.

Contemporary CATV networks are of the hybrid fiber-coax (HFC) type. Their digital signals transmission is based on the DVB-C standard. Thus, signals transmission over the downstream channel is carried out through quadrature amplitude modulation (64QAM or 256QAM, as usual) of the radiofrequency (RF) carriers within the RF band from 450 MHz to 862 MHz.

Information coming from the subscribers is transmitted over the upstream channels within the band from 5 MHz to 65 MHz, noise-immunity methods (mostly QPSK and 16QAM) being used to this end. The downstream channel's width is 7 MHz or 8 MHz (the European standard) while the upstream channel's width varies from 200 kHz to 6,4 MHz depending on the channel capacity requested. The digital signals' transmission bit rate over the downstream channel can be up to 50 Mbps if additional information processing methods (compression and coding) and 256QAM are applied. As for the upstream channel with 16QAM the bit rate is about 9 Mbps.

According to references [1-5], different methods for solving the problem are used, which are based on the optimization of the CATV network architecture to provide VoD service and on the movie content distribution algorithms as well. Now-a-days methods are usually based on the Peer-to-Peer protocol in the subscriber's side of the network. Most of the suggested solutions however turn out to be impracticable for CATV networks because of the limited upstream bit rate (for upstream channels with QAM-16 and channel width of 6 MHz the bit rate is about 9 Mbps).

Besides, with the tree-and-branch network topology no direct data transfer between two different subscribers is possible, the transfer is only made via the higher hierarchical level of the network (distribution hub or regional head-end). The subscribers' set-top-box devices are with limited storage capacity, so they should be numerous enough (large population networks) in order to effectively apply the Peer-to-Peer (P2P) technology in VoD service distribution. Hence, optimal solutions are needed with respect to the traffic across the central transport network, the servers' capacity and the equipment cost.

The VoD systems must be designed in a way to achieve optimum video traffic distribution over the network in order to obtain maximum effectiveness and minimum expenses. In order to determine the distribution laws of various random events long-term investigations are needed based on preliminary observations and statistical data processing.

In [6] the investigation results referring to the following random events are discussed: the video traffic within a 24-hour period, the time between incoming requests for VoD service delivery in the most-watched hours, the duration of the video streams requested, the movies distribution according to subscribers' preference.

The main purpose of the paper is to study the possibilities to optimize the VoD system in CATV networks in terms of the edge servers' capacity and to develop an algorithm for automatic movie content distribution in order to minimize the traffic across the central transport network, hence to minimize the blocking probability in the system.

Manuscript received February 5, 2010
Manuscript revised February 20, 2010

2. Architecture and operation algorithm of the VoD system

Modern VoD systems realized over CATV networks are based on the hierarchical approach, the number of hierarchical levels being usually three. The edge VoD servers located on the third hierarchical level hold the movies most frequently requested by the corresponding subscribers' group. The video stream transfer is carried out from those servers to subscribers. The servers can function as proxy servers as well to redirect the video stream from another server, which holds the movie requested by the subscriber. The VoD servers with saved movies that are most-requested in the corresponding region are located on the second hierarchical level. Between them and the edge VoD servers information is periodically exchanged, containing the number of requests for the corresponding region. The number of requests is different for each separate group. The edge nodes are updated periodically in accordance with the subscribers' preference. On the first hierarchical level the head-end to coordinate the system performance is located.

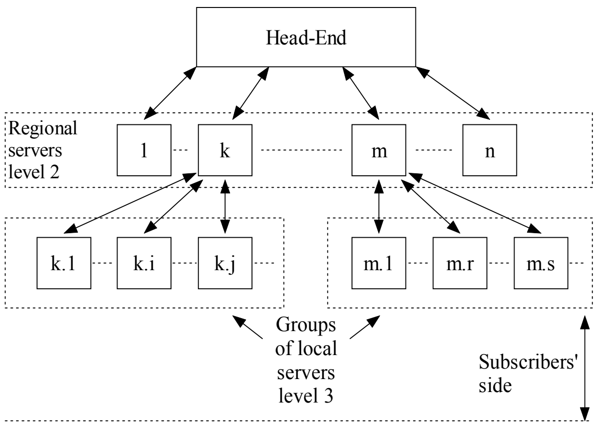

Fig. 1 VoD system architecture.

The VoD system architecture here discussed is shown in Fig. 1. Regional servers numbered 1 to *n* are installed within the network during the construction, their number being equal to the number of local servers group on level three. With the system deployment additional servers on level 3 with theirs served subscribers are installed.

There are two algorithms (direct and indirect) for the VoD system operation. When the requested movie is in the particular subscriber's region, the direct algorithm is used. This algorithm works as follows: The movie request of the subscriber is sent to its local server on hierarchical level 3 (assume it is numbered *k.i*) that forwards it to its manager (server *k* in the case). When server *k* finds out that a copy of the requested movie is available in some of the local servers from the *k*-th group (let it be server *k.j*) and the video transfer is possible, then server *k* replies to server *k.i*, the latter in its turn replying to the subscriber with information about the number of the server that contains the requested movie. At the same time server *k.j* informs server *k* about the request expected from the user. Then, for a certain period of time server *k.j* reserves the necessary resources to respond to the request. Eventually the subscriber sends a request to server *k.j* giving information about the movie title and VCR commands. Then the video stream transfer from this server starts.

The indirect algorithm is used either when the requested movie is stored outside the local region or when all local servers that hold a copy are overloaded. Though server *k.j* which holds the requested movie is overloaded, it can still serve as a proxy server. In this case, the manager (in the case server *k*) forwards the request to the head-end. The head-end starts searching a regional server where there is a copy of the movie and the video stream transfer to the subscriber is possible via server *k.j*. The regional server chosen (let it be server *m*) will look for a local server that holds a copy of that movie, is able to respond to the request and is connected to server *k.j*. If such a local server is detected (let it be server *m.r*) the subscriber will contact it to give his request, then the transfer of the requested video stream via server *k.j* will start.

3. Determination of the average duration and file size of the movies in the system

The investigation is based on statistical data about the duration *L* and file size *S* of 1100 popular movies (550 in the SD and 550 in the HD format). Information supplied by a cable operator with 100 serving centers (each one of about 200 subscribers) was collected during a 6-month period. To determine the distribution law of the random values *L* and *S* the MS Excel Data Analysis Module was used.

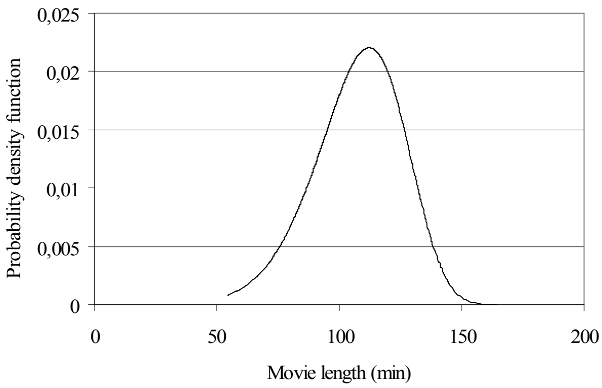

Fig. 2 Movie length distribution.

In Fig. 2 the distribution law of the movie duration L is shown. It is obvious that the distribution of the random value under research may be approximated accurately enough by a Gaussian law and described with the following expression:

$$f(L) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(L-M)^2}{2\sigma^2}\right] = 0,02493 \exp\left[-\frac{(L-113,8)^2}{512}\right] \quad (1)$$

where M (113,8 min) is the mathematical expectation value and σ (16 min) is the standard deviation of random L , the boundary values of L being $L_{\min} = 54$ min and $L_{\max} = 164$ min, respectively.

During the investigation of the file size distribution S the transmission bit rate was assumed to be 3,75 Mbps for the SD and 15 Mbps for the HD movie format. The results from statistical data processing of the measured random values show that the distribution of S can be described by a normal law (Fig. 3).

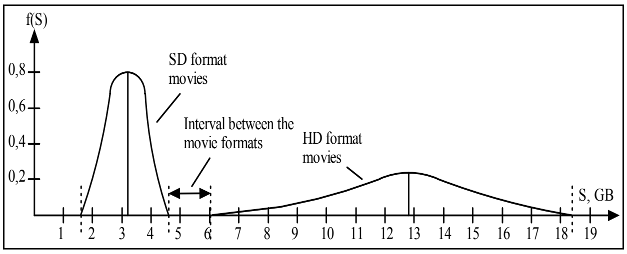

Fig. 3 File size distribution laws for movies in the SD and HD format.

The distribution law parameters such as variation boundaries S_{min} and S_{max}, mean value S_0 and standard deviation σ are given in Table 1.

Table 1: File size distribution law parameters

Parameter	$S_{\min}$	$S_{\max}$	S_0	σ
Movie size SD, GB	1,52	4,61	3,2	0,45
Movie size HD, GB	6,08	18,45	12,8	1,8

The movies' file size measured for both formats makes it possible for the movie type to be defined. If the file size value S is from 1,52 to 4,61 GB, it follows that the transmitted movie is in the SD format. A value of S in the interval from 6,08 to 18,45 GB denotes the HD format. These results can be used to develop an algorithm for the movie content distribution that is used to determine the movie type by its file size. An important condition for the determination of SD and HD file sizes is that they must not intersect. In the case this condition is fulfilled.

4. Results of research on movie content popularity

Movies provided through VoD systems can be classified into the following three categories: mass movies, other movies and VoD browse. Mass movies can be divided into two categories – series with a large number of episodes and one-episode movies. The paper deals with investigation carried out to study the drop in video series popularity observed at a certain moment after the series' broadcast. To this end statistical data have been collected about series, broadcasted from a CATV operator with 100 local centers, each of them with 250 subscribers. In Fig. 4 the distribution law of the average number of video sessions is shown within a period of two weeks after the series have been broadcasted.

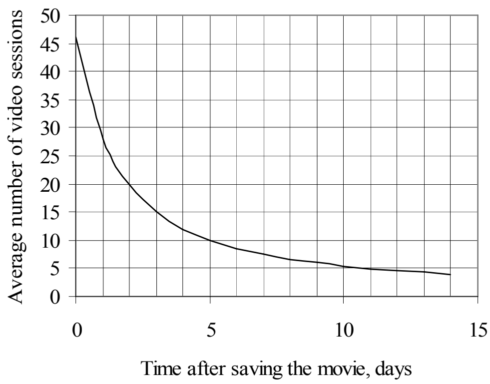

Fig. 4 Decrease in video series popularity.

The results of the research can be used to determine the time period during which series must be kept stored onto local servers. If the time period to store series in the server and the average file size S_0 are known (value S_0 is assumed to be the same for movies in the SD format) the amount of storage capacity to be set apart for series stored in each of the local servers can be easily determined. This is of particular importance for the appropriate design of the algorithm for movies' content distribution in VoD systems.

3. On the network optimization in terms of local servers' capacity

A fundamental problem with VoD systems is the traffic across the central transport network which is related with the probability of blocking the system. This probability can lessen if the movies are appropriately distributed among the local servers. In Fig. 5 the possible change in the local servers' capacity C_{ls} is shown (assuming that all local servers have the same capacity). In the case, $C_{\max}$ is the maximum local server's capacity (if assumed that it does not change for a long period of time); C_{opt} is the optimum capacity denoting the boundary between static and dynamic movie content in the server. If $C < C_{opt}$, the central transport network will be overloaded with video traffic. If $C > C_{opt}$, the subscriber is to wait frequently enough for the delivery of the requested movie. The value C_{opt} is determined by the system itself through local servers' refresh, performed once in a 24-hour period in the time interval between 4 AM and 6 AM when the video traffic is at its minimum.

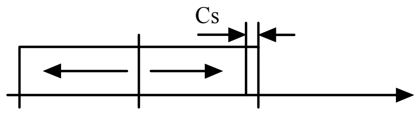

Fig. 5 Local server capacity distribution.

The maximum local server's capacity C_{max} is chosen to be at least ¼ of the regional server's capacity. Such a value will guarantee that 80 % of the movies requested in the respective service area will be available at the local server. This follows from Zipf's (Pareto's) law for subscriber's preferences distribution which is known in literature as the "80/20 rule". The higher the value C_{max} the better the system performance, though this can cause a dramatic increase in local servers' cost. Therefore, the determination of C_{max} requires the cable operator to make a compromise between cost and quality.

Since the local server's capacity is limited, the saved amount of movie content must reply to the following condition:

$$\sum_{i=1}^{p} S_{i\,SD} + \sum_{j=1}^{q} S_{j\,HD} + C_S = C_{\max}, \tag{2}$$

where $S_{i\,SD}$ and $S_{j\,HD}$ are the file sizes of the i-th movie in the SD format and the j-th movie in the HD format respectively, p and q are the numbers of movies in the SD and HD format, which are saved onto the local server (without the number of saved series in the SD format), and C_S is the local server's capacity set aside in reserve for series.

The values of parameters $S_{i\,SD}$ и $S_{j\,HD}$ can be easily determined if the file size of a movie of one-hour duration S(1h) is known. It can be determined in GB with the following expression:

$$S_{(1h)} = 0{,}45.R, \tag{3}$$

where R is the video stream bit rate in Mbps. The authenticity of the upper expression is also confirmed by the online calculator in [7].

The bit rate of MPEG-2 TS movie transmission being 3,75 Mbps in the SD format, and 15 Mbps in the HD format respectively, the following expressions for the movie file size determination (in GB) in both formats can be obtained from (3):

$$S_{i\,SD} = 1{,}6875 L_i \qquad S_{j\,HD} = 6{,}75 L_j, \tag{4}$$

where L_i and L_j are the duration (in hours) of the movie with the corresponding number.

The peculiarity of transmitting series with a large number of episodes (usually more than 100) that are daily broadcasted in the SD format (as a rule) must be taken into consideration. As shown in Fig. 4, the popularity of such a kind of movie is at its maximum right after the broadcast and drops almost to zero within two weeks. Because of that it is assumed that the series are saved onto all local servers immediately after the first broadcast and are automatically deleted after some period of time in order to eliminate the transit traffic referring to that type of movies. With such an approach the disk storage losses are negligibly low if related to system performance improvement (in terms of the transit traffic). Practically, all series in one-week period can be saved onto disk storage of some tens of Gbytes, the storage capacity of contemporary VoD servers [8] amounting to several tens of Tbytes, i.e. nearly 1000 times larger.

3. Algorithm for optimum distribution of the video content

In the paper an interpolation algorithm is considered that distributes movie content onto the local servers according to the movie index variation. This algorithm is shown in Fig. 6 and represents an improved version of the algorithm proposed in [6].

In order to make a decision about the allocation of a movie the following parameters must be considered: free disk space on each of the local video servers, the number of requests made for session time-out values over 20 minutes, number of unattended requests to subscribers. Besides, long TV series (over 100 episodes) are eliminated from the indexing system as they are copied in all local servers immediately after the broadcast and are automatically deleted after a certain period of time which is optional for the operator (about 3 to 10 days after the broadcast). These series should be marked out by the system in order not to be indexed and distributed by the algorithm (as the other movies are).

The probability to select the n-th most popular movie out of M films in the system can be determined by the following expression [9]:

$$P_n = \left(n \sum_{k=1}^{M} \frac{1}{k} \right)^{-1}, \tag{5}$$

where k is the sequential number of the movie.

Thus, operation of the algorithm is reduced to loading and deleting the movie content in local servers at a daily basis according to given criteria and at a precisely defined time-interval. The interval from 4 AM to 6 AM when video traffic is at its minimum turns out to be most suitable for the purpose. Two parameters are used as criteria for entry or deletion of the movie content: 1) index $I(k)$ of the k-th film indicating the number of video sessions hold for that film and 2) parameter $R(k)$ which denotes the number of outstanding requests for the k-th film. An outstanding request means inability to deliver the requested movie to the subscriber rather than waiting for the download of the selected movie from other servers when it is not stored in the local server. Parameter $R(k)$ is used by the algorithm to determine the movies in the relevant video server with the largest number of outstanding requests and to store these movies into the adjacent local servers when the video content is refreshed.

The suggested algorithm is described as interpolation, since the determination of both the movie index $I(k)$ and parameter $R(k)$ is based on conjectural values that may be attained during the peak hours in terms of video traffic i.e. about 8 PM to 10 PM [6]. Or, in other words,

$$I(k) = I(k) + dI(k)$$
$$R(k) = R(k) + dR(k)\,, \tag{6}$$

where $dI(k)$ and $dR(k)$ are the differences between the current value of parameters $I(k)$ and $R(k)$ respectively and the values these parameters would attain during the time period from 8 PM to 10 PM when video traffic is at its maximum.

Parameters $I(k)$ and $R(k)$ are recorded in the database immediately after the refresh procedure is activated. The database can be installed in the head-end and/or in the regional VoD servers. Based on it the average values of $I(k)$ and $R(k)$ for the last few days preceding the current refresh are calculated and used to obtain the values of $dI(k)$ and $dR(k)$ that the algorithm needs.

To compensate for the difference in the transmission bit rate of the movies in the HD and SD format it is necessary to multiply by 4 the index of the movie in the HD format, the decision being taken after the movie file size $S(k)$ is considered. After completion of the refresh procedure values $I(k)$ and $R(k)$ are set to zero.

The procedure of refreshing the video content consists of gradually deleting the movies from the local VoD servers and recording new movies in accordance with their conjectural index and number of unattended requests.

Unlike the algorithm in [10] where calculation of the required number of the movie copies is based on the number of unattended requests for that movie, the algorithm here described involves a stage-by-stage procedure. The movie index originally based on conjecture is compared with a given threshold value. If the threshold is exceeded the movie is recorded onto the local server. Then the conjectural number of unattended requests to this movie is compared with another threshold value and if it is still above it a copy of the movie onto the adjacent server is made so that it can be transferred from there to the subscriber, the local server operating as a proxy one. The procedure here described results in the following formula:

$$p\sum_{x=1}^{N_p} C_x + q\sum_{y=1}^{N_q} C_y = C_{opt} \tag{7}$$
$$p + q = 1\,,$$

where p and q are threshold coefficients to determine the weights of parameters $I(k)$ and $R(k)$ when making a decision to record the corresponding movie; C_x and C_y are the disk space occupied by the corresponding movies; N_p and N_q are the number of movies that should be recorded onto the local VoD server if parameters $I(k)$ and $R(k)$ are taken into consideration.

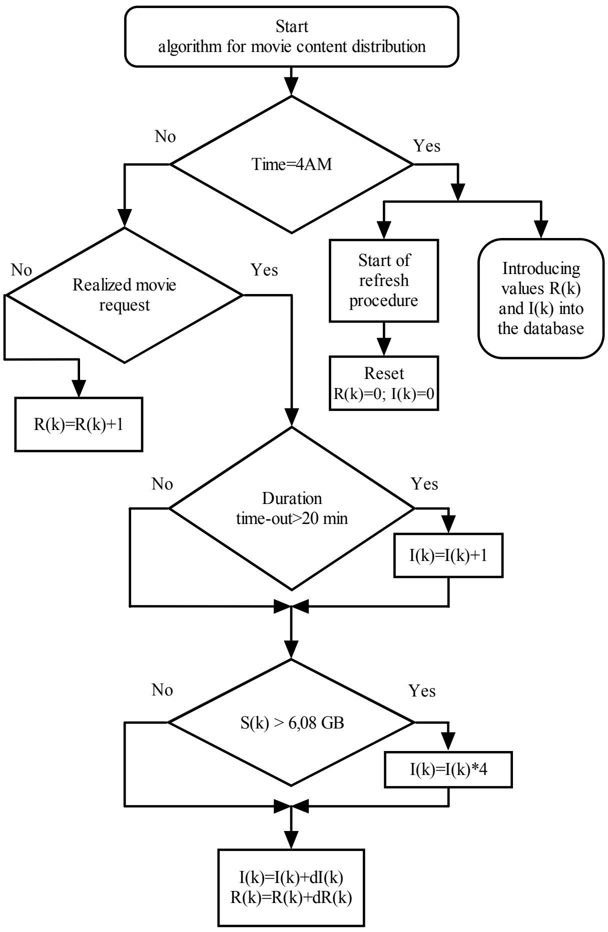

Fig. 6 Interpolation algorithm for movie content distribution.

The process of recording the movies onto the local server continues until the server optimal (threshold) capacity C_{opt} is attained. The value C_{opt} is determined by the algorithm itself. To this end information is collected about the ratio of the number of video sessions to movies recorded into the local server and the number of video sessions to movies downloaded from other servers. Information about this ratio is stored in the database at the beginning of the renewal process, the threshold C_{opt} being shifted with a given step at each update until the ratio gets its maximum. The optimization of the VoD system in terms of server capacity is over when the value of C_{opt} is determined for each of the local servers. The process can last from several days to several weeks.

3. Conclusion

On the basis of the interpolation algorithm for movie content distribution suggested in the paper the following conclusions can be drawn:

1) As for the video content the trend is that 80% of the subscribers would prefer about 23% of the video content offered.
2) The movie duration and file size follow the normal distribution law.
3) The type (SD or HD) of the movie can be simply determined by the file size.

The decrease in the series' popularity can be described more exactly by the Gamma distribution, the maximum value of the movie's popularity being obtained right after its broadcast and saving into the system. Two weeks later the popularity drops to the zero.

References

[1] Allen M. S., Zhao B. Y. and Wolski R. Deploying Video-on-Demand Services on Cable Networks, *27th IEEE International Conference on Distributed* Computing Systems (ICDCS 2007), 2007.
[2] Gong J., Reed D., Shaw T., Vivianco D. and Martin J. QAM Resource Allocation in mixed-Format VoD Systems, Jan. 2007.
[3] Understanding User Behavior in Large-Scale Video-on-Demand Systems. Proc. First ACM SIGOPS/EuroSys European Conf. Computer Systems (EuroSys '06), Apr. 2006.
[4] Yinqing Zhao Kuo, C.-C.J. Scheduling Design for Distributed Video-on-Demand Servers. Dept. of Electr. Eng., Univ. of Southern California, Los Angeles, CA, USA, 2005.
[5] Huang Y. and Venkatasubramanian N. Data Placement in Intermittently Available Environments. International conference on high performance computing, N^{o}9, Bangalore, INDE (18/12/2002).
[6] Jordanova L., Nenkov J. Distribution of Video-on-Demand Service over Cable Television Networks. Radioengineering, Vol. 2, No 09-021, 2009.
[7] http://www.numion.com/calculators/units.html
[8] http://www.thomson.net/SiteCollectionDocuments/Solutions/Data%20Sheets/Datasheet_Thomson_SmartVision_VOD_Apr09.pdf
[9] Tsang, K. and Kwok, Sai. Video Management in Commercial Distributed Video on Demand (VoD) Systems. PACIS 2000 Proceedings, paper 17.
[10] Wei-tao X., Peng L., Yu-ping M., Zhao-hong Y. Collaborative Replica Selection and Substitution Algorithm for the Video Grid. Journal of Communication and Computer, Sep. 2005, Volume 2, No.9 (Serial No.10).

Lidia Jordanova received M.S. and Ph.D. degrees in radiocommunications technologies from the Technical University of Sofia, *Bulgaria*, in 1972 and 1978 respectively. From 1980 to 1987 she was a research engineer with the Institute of Radioelectronics in Sofia, Bulgaria, where she was engaged in antennas and microwave circuits design. She is presently an Assoc. Professor of Telecommunications Equipment Design, and Head of the Center for wideband communications and CATV at the Technical University of Sofia. She is the author of more than 100 publications, 11 books and 4 BG patents of invention and has been engaged in more than 60 projects in radio relay and CATV systems and modules. Her current research interests are in HFC CATV systems and fiber-optics integrated circuits.

Jordan NENKOV was born in 1977. He received M.S. degree in aviation electrical engineering from the National Military University – Aviation Faculty, Dolna Mitropolia, Bulgaria in 2001. He is presently Ph.D. student at the Technical University – Faculty of Telecommunications, Sofia, Bulgaria. His current research interests are in VoD and IP-based services over HFC CATV systems.