How to cite this paper:

Suliman Mohamed Fati, Putra Sumari, Siti Sophiayati Yuhaniz, & Nilam Nur Binti Amir Sjarif. (2017). Modelling contents status for IPTV delivery networks in Zulikha, J. & N. H. Zakaria (Eds.), Proceedings of the 6th International Conference of Computing & Informatics (pp 282-290). Sintok: School of Computing.

MODELLING CONTENTS STATUS FOR IPTV DELIVERY NETWORKS

Suliman Mohamed Fati[1],Putra Sumari[2], Siti Sophiayati Yuhaniz[3], Nilam Nur Binti Amir Sjarif[4]

[1]*Faculty of Information Technology-Math & Science, Inti Internatioanl University, Persiaran Perdana BBN, Putra Nilai, 71800 Nilai, Negeri Sembilan*
[2] *Multimedia Research Group, School of computer sciences, Univeristi Sains Malaysia, Malaysia*
[3,4]*Advanced Informatics School, Universiti Teknologi Malaysia,54100 Kuala Lumpur, Malaysia.*
[1]*smfati@yahoo.com,* [2]*putras@usm.my,* [3]*sophia@utm.my,*[4]*nilamnur@utm.my*

ABSTRACT. Since IPTV has been invented, IPTV is considered a dominant technology to distribute high quality videos and live channels anytime anywhere over challenging environment to end users who are having different preferences and demands. Presently, IPTV service providers manage IPTV delivery networks, in terms of contents, channels, resources, based on contents popularity distribution and/or users' preferences only. Although content popularity and users' preferences play an important role to cope with the increasing demand of IPTV contents/channels, these two measures fail in producing efficient IPTV delivery networks. For that, IPTV delivery network designing should integrate the IPTV content characteristics like size, interactivity, the rapid changing lifetime. Therefore, the idea of this paper is to build a mathematical model that integrates all these factors in one concept called IPTV content status. Modeling the contents status according to its characteristics is an important point to design Content-Aware IPTV delivery networks. The experimental results showed the superiority of modeling IPTV content status in balancing the load and reducing the resources waste.

Keywords: IPTV, Delivery Networks, Content Popularity, Contents' workload, IPTV content status.

INTRODUCTION

With the massive development in networking and multimedia fields, IPTV has become recently popular as a promising trend in home and business entertainment due to the increasing number of IP-networks users. Therefore, IPTV, as a promising technology, grows rapidly in terms of subscribers and revenue to become, in the near future, the standard means to deliver home and business entertainment contents (Yarali & Cherry, 2005). Thus, Telecommunication companies have entered a hectic competition to increase their customer base and profit by providing IPTV services (Yarali & Cherry, 2005) (Lee, et al, 2009) (Li & Wu, 2010). Nowadays, IPTV is globally considered a dominant technology to distribute high quality videos and live channels anytime anywhere over challenging environment. This challenging environment tends to deliver a hybrid IPTV services over IP-based network infrastructure into end-users with different preferences and demands (Aldana Diaz & Huh, 2011) (Montpetit, Klym & Blain, 2010) (Montpetit, Klym, & Mirlacher, 2011).

Currently, IPTV service providers consider the IPTV contents popularity distribution and/or users' preferences to allocate IPTV contents (Dukes & Jones, 2004) (Sujatha, et al, 2008) (Lie, Lui, & Golubchik, 2000) (Little & Venkatesh, 1993) (Sobe, Elmenreich, & Böszörmenyi, 2010), balance the network workload (Huang & Fang, 2004) (Zhou & Xu, 2002), and/or pre-fetch the desired channels into the Set-top-box (Bikfalvi, 2011). Moreover, IPTV service providers exploit the users' preferences and content popularity to recommend some contents and also to control the advertisement that should be displayed to the users.

Although content popularity and users' preferences play an important role to cope with the increasing demand of IPTV contents/channels, these two measures fail in producing efficient IPTV delivery networks (Fati, Sumari, & Budiartu, 2015) (Fati, Budiartu, & Sumari, 2014) (Fati & Sumari, 2016). For example, allocating the IPTV contents according to the content popularity only may lead to satisfy the users' demand or preferences. According to (Joe, Yi & Sohn, 2011), the popularity-based replication schemes assume a threshold to decide which contents should be replicated. In other words, the contents having popularity values larger than this threshold can be replicated. Also, some of these algorithms are proactive wherein the contents can be replicated once the server allocating these contents becomes overloaded. Actually, replicating the huge size contents, like IPTV contents, should consider the trade-off between the QoS requirement and exploited storage space. In other words, the replication algorithms should maintain the QoS requirements without wasting the storage resources. Both the popularity based and proactive algorithms may replicate redundant contents, which leads in turn, to increase the exploited storage space. On the other hand, the current studies focus on the channel popularity and/or user preferences only. In the real world, the popularity of channel or the user behaviour suffer from the rapid fluctuation. Such fluctuation in channel popularity and/or users' preferences may lead to prefetching irrelevant channels, which in turn, leads to bandwidth saturation and delay the required channel, as well.

To overcome this challenge, the design process of IPTV delivery networks for both On-demand and live services should consider the other characteristics of IPTV contents/channels. Beside the popularity, these characteristics include the huge size (e.g. HD requires huge bandwidth more than the normal channels), interactivity, the rapidly changing effective (Joe, Yi & Sohn, 2011). In addition, there are other factors should be considered including multicast factor (e.g. some channels are requested by many people in the area while other channels requested rarely or by individuals. The population size is another factor should be considered (e.g. some service areas are crowded and then the request rate is higher than other areas.

According to (García, et al, 2009) (Fati & Sumari, 2016), modeling the workload for multimedia streaming services is essential factor in delivery networks to improve the performance, enhance the QoS, and increase the reliability. Therefore, the idea of this paper is to build a mathematical model that integrates all these factors in one concept called IPTV content status. Modeling the contents status according to its characteristics is an important point to design Content-Aware IPTV delivery networks. Content status modeling refers as estimating the portion of concurrent requests that target the content according to its characteristics. Such modeling helps a lot in handling both the load balancing and resource allocation based on the anticipated load of contents. However, these contents are huge size with non-uniform patterns of both contents lifetime and users request access. For this reason, modeling the contents status is a challenge. Many works in the literature review focus on the popularity of contents during their approaches to solve load imbalance problem. However, content popularity is not everything. The other characteristics like the size and the interactivity play a significant role in modeling the content status. Hence, the aim of this paper is how to model the content status in IPTV environment. Modeling the content status helps the delivery networks providers to build content-aware delivery networks to overcome the load imbalance and resources waste problems. To the best of our knowledge, there is no work formulate the content status for IPTV contents.

BACKGROUND AND RELATED WORK

The expected load of content or server has a significant role in balancing the load among the servers. Estimating the load of content or server takes different forms according to the nature of problems. In some works, the load can be estimated as the CPU consumption while, in other works, the load can be estimated as the active connections to the servers (Cho, et al, 2008) (Meng, Liu, & Yin, 2013) (Kanrar, 2012). In the present paper, the load is estimated as the number of active concurrent requests that access the content/servers. Thus, the IPTV content status model that estimates the maximum load for each video in the service area at the peak busy period is suggested. The video load is estimated based on its popularity, length, and request arrival rate (e.g. normal requests and interactive requests) (Gaber & Sumari, 2014) (Fati, Budiartu, & Sumari, 2014). The proposed estimation model deals with both the normal and interactive viewing sessions. The load and the workload will be used interchangeably throughout this thesis.

For each video, the workload can be estimated as the number of concurrent requests that must be served by that video. The estimated workload for a video is a function of the request arrival rate, the number of subscribers within the service area, the video popularity, and the video length. All these information, which are required in the workload estimation, should be stored in the request dispatcher that is located in front of the clustered VHO. To estimate the workload in IPTV system, we have to imagine the viewing scenario of any subscriber as follows: Once connected, the subscriber has two viewing states; a normal state and an interactive state. The user starts in the normal state where the requested video is being viewed at the usual speed. After that, the subscriber issues an interactive VCR command to stop, pause, and jump forward or backward (Cho, et al, 2008). During the viewing time, the user can transit from one state to another interchangeably until the end of viewing time.

Based on the above viewing scenario, there are two types of requests: normal requests and interactive requests. Thus, the expected workload of a video i can be defined as the portion of concurrent requests of both types that can be served by this content within the peak busy period. Therefore, the expected load of content i is estimated as the fraction of simultaneous incoming requests targeting content i.

According to (Kanrar, 2012) (Wong, 2004), the number of concurrent requests at a certain time point can be estimated as the sum of the active requests lengths within an interval time divided by this interval time. Kanrar (Kanrar, 2012) has estimated the server's workload as a function of number of households, request arrival rate, and video length during the peak busy period.

Most of studies focus on the channel popularity and/or user preferences only without any consideration for the channel characteristics like size (e.g. HD requires huge bandwidth more than the normal channels), multicast factor (e.g. some channels are requested by many people in the area while other channels requested rarely or by individuals, and the population size (e.g. some service areas are crowded and then the request rate is higher than other areas. Moreover, the popularity of channel or the user behavior IPTV systems suffer from the rapid fluctuation of users' preferences in the real world. Therefore, such fluctuation in channel popularity and/or users' preferences may lead to prefetching irrelevant channels, which in turn, leads to bandwidth saturation and delay the required channel, as well. For that, expecting the channel status (the expected requests targeting this channel) according to the channel characteristics and demand is a beneficial. This prediction can control the fluctuation in the channel popularity and/or the users' preferences. The channel status can be modeled as a function of channel bandwidth, channel popularity, multicast factor, and region population. According to the history analysis, we can easily, predict the demand of this channel in a particular area, then, decide the prefetching.

IPTV CONTENT STATUS MODEL

As mentioned above, the workload of contents can be estimated by considering the content's characteristic. For example, the content workload means the number of active requests targeting that content at the same time during the peak busy period. In the case of estimating the workload for individual content, there is a population of subscribers H issuing different requests targeting diverse contents at the peak busy period. Each subscriber can issue a number of normal requests λ and interactive request λ_{vcr} with holding times t_i and t_{vcr} for normal and interactive requests, respectively. Thus, the workload for a video i can be estimated as the portion of concurrent normal and interactive requests originated within the service area by the subscribers to watch that content during the peak busy period. This can be interpreted mathematically as in Eq. (1).

$$L_i = \frac{p_i * H}{T_{peak}}((\lambda * t_i) + (\lambda_{vcr} * t_{vcr})) \quad (1)$$

Where the terms L_i, p_i and t_i denote the expected work load, the popularity and the length of the particualr movie i in the service area, respectively. The popularity takes a value between 0 and 1. The term H denotes the number of households in the service area, and T_{peak} denotes the peak busy period of IPTV system in minutes. In addition, the terms λ and λ_{vcr} denote the request arrival rate and VCR commands request arrival rate, respectively. Both arrival rates follow Poisson distribution. In equation (1), the term $H((\lambda * t_i) + (\lambda_{vcr} * t_{vcr}))$ represents the total sum of all requests issued by the household s within the peak busy period. By multiplying this term by the content popularity value p_i, the number of requests targeting this content is obtained. Finally, the concurrent requests per a unit time are estimated by dividing on the term T_{peak}.

After contents workload estimation, the expected workload of server j is estimated by summing up the load of all contents that are stored in this server as shown in Eq. (2).

$$L_j = \sum_{i\in j} L_i = \frac{H}{T_{peak}} \sum_{i\in j}(p_i((\lambda * t_i) + (\lambda_{vcr} * t_{vcr}))) \quad (2)$$

The term L_j denotes the expected workload for a particular server j. Assuming that the available servers in the service area are sufficient to handle the total workload, thus, these servers must be allocated accurately to cope with the growing load problem. After estimating the expected load of contents, the replication degree (i.e. the number of replicas that are required to handle the expected load) is calculated. The replication degree must be controlled by the video popularity. For instance, video i can be allocated on all servers if its popularity is extremely high. This is to ensure that the expected load can be distributed on as a large number of servers as possible to minimize the request rejection rate. On the other hand, there is no need to replicate the remarkably low popular videos so that one copy is enough to catch its low expected load. Based on the above explanation, the replication degree (i.e. number of copies) for a video i can be formulated as a function of the number of servers and the normalized popularity as shown in Eq. (3). According to (Nafaa, Murphy & Murphy, 2008), normalizing the popularity distribution improve the overall performance by building a strong relationship between the content popularity and the replication degree.

$$r_i = \lceil S_a * p_i^n \rceil \quad (3)$$

Where the term r_i denotes the number of replicas for content i , S_a represents the number of servers in service area a , and finally the term p_i^n representat the normalized value of p_i that should be within [0,1]. The normlized popularity value has been obtained using Min-Max Normalization Law. Min-max normalization performs a linear transformation on the original data (Gaber & Sumari, 2014). Min-max normalization maps p_i to p_i^n value in the range $[\text{new\_min}, \text{new\_max}]$. In this case, the new_max value equals one where the highest popularity value in the dataset will be scaled to equal the value (1) and the other values will be scaled successively. The operator $\lceil . \rceil$ is a ceiling function operator to take the largest integer nearest to the calculated term.

After computing the expected number of replicas for a video i, the expected load for each replica can be calculated by dividing the expected load for that video on its number of replicas as follows:

$$L_{ri} = L_i / r_i \tag{4}$$

Where L_{ri} represents the load of one replica for content i, L_i denotes the load of video I, and r_i represents the number of replicas for content i, which obtained from Eq. (3).

According to the proposed content status model, the content of high popularity value, which has a highly expected load, is replicated more to handle the increasing demands. On the other hand, the contents of low popularity value can be replicated as less as possible for serving the low expected load. Furthermore, allocating the contents and distributing the incoming requests according to this proposed IPTV content status model can be useful in maintaining the load as balanced as possible within the service area.

EXPERIMENTAL RESULTS

To investigate the performance of our proposed IPTV content status model, we have tested the model on an empirical data set that is sampled according to the content popularity distribution. The popularity distributions used are Zipf's like distribution (Zipf). There are a set of assumptions that taken into account during the experimental study of the proposed model as following: the population of subscribers H=1000, each subscriber can issue a number of normal requests λ=2 request/user/minutes and interactive request λ_vcr=2 request/user/minutes with holding times t_vcr=10 seconds for interactive requests. These values are typical empirical data and widely used by different studies (Cho, et al, 2008). The content size distribution has been generated randomly between 50-1000 MB. Figure 1 depicts a sample of the content size distribution in MB.

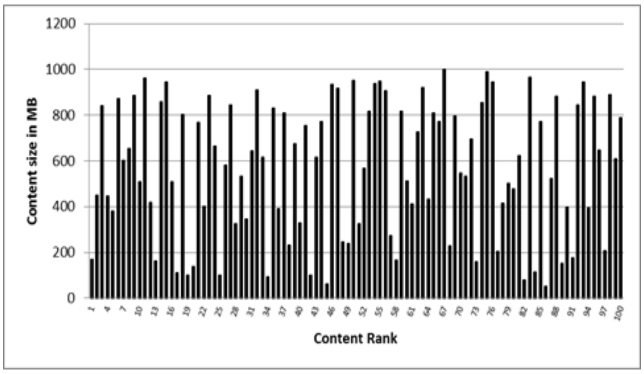

Figure 1: Random Contents size Distribution

The popularity of contents is computed according to Zipf's Law (Li & Wu, 2010) as in Eq. (5) where i, N, and θ represent the content rank, the total number of contents, and the skewness degree respectively. The content rank refers to the position in of content in the sorted contents list. The sorted contents list is sorted according to their popularity values. In this list, the first video is the highest popular video, and so on.

$$P(i, \theta, N) = \frac{1/_{i^{\theta}}}{\sum_{n=1}^{N}\left(1/_{n^{\theta}}\right)} \tag{5}$$

Then, the workload is computed using our proposed IPTV content status model. Note that the expected load of server j can be expressed by summing the load of all contents stored in this server. Figure 2 depicts the expected load for the contents, which is obtained from the IPTV content status model.

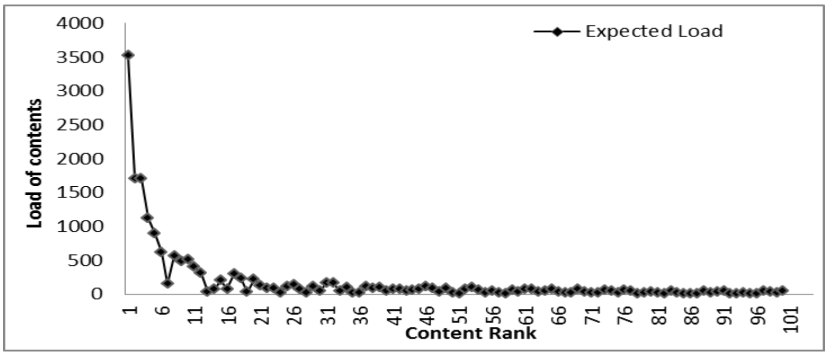

Figure 2: The distribution of expected content load

Figure 3 plots the relationship between the expected workload of contents and the popularity distribution. What is apparent is the fact that the expected load of contents differs from the popularity distribution. This is due to the other factors that contribute in the workload estimation like video length and interactivity.

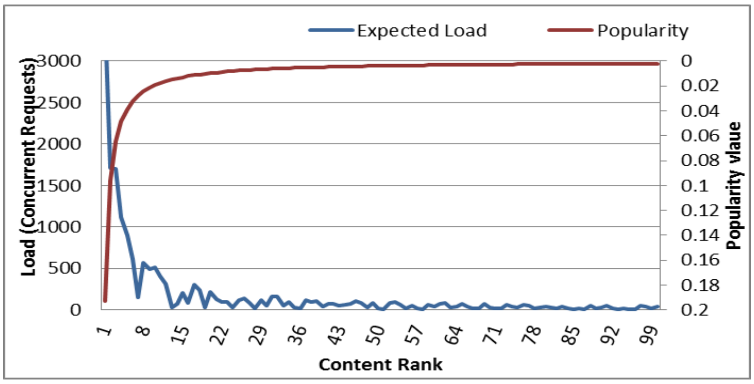

Figure 3: The relation between expected load and popularity

Another experiment is conducted on a delivery network with 4 service areas. Each service area contains four servers. We changed the popularity skewness at each service area such that each area allocates different replicas for the contents. After that, our request distribution algorithm (Gaber & Sumari, 2014) has been developed over the IPTV content status model. Then, this algorithm (CARDA) has been compared with other two request distribution algorithms, Chu (Cho, et al, 2008) and Round Robin (RR) algorithm. The experiment showed that the popularity has no effect on the request distribution mechanism that supports the argument in (Huang, et al, 2005). Chengdu (Huang, et al, 2005) argued that the request distribution process is solely affected by the request arrival rates. Figure 4 depicts the distribution of workload among four service areas with four servers for each. Moreover, it is apparent the facts that our model, which exploit the content status, has a balanced workload distribution among all servers.

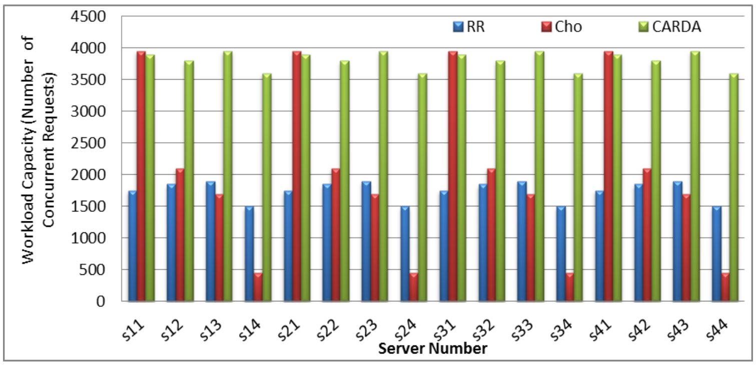

Figure 4: The relation between expected load and popularity

CONCLUSION

Managing IPTV delivery networks that allocate interactive huge size contents (video and/or channels) characterized by fluctuated popularity and non-uniform request patterns is challengeable. Ignoring such characteristics lead to load imbalance and affects the system performance. Therefore, modelling the contents status according to its characteristics is very important in designing IPTV delivery networks. Such modelling helps a lot in handling the load balancing by distributing the incoming requests based on the anticipated load of both contents and servers. However, modelling the contents status is challengeable due to the nature of the IPTV contents. Most of the proposed works focus only on content popularity, but the influence of other characteristics should be investigated. Thus, this thesis suggested the IPTV content status model that estimates the maximum workload for each video. The load is estimated as the number of active concurrent requests that access the content. The workload for the content is estimated in the service area at the peak busy period based on its popularity, length, and request arrival rate (e.g. normal requests and interactive requests). The proposed model deals with both the normal and interactive viewing sessions.

REFERENCES

Aldana Diaz, M. E., & Huh, E. N. (2011, February). Cost analysis on IPTV hosting service for 3rd party providers. *In Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication* (p. 114). ACM.

Montpetit, M. J., Klym, N., & Blain, E. (2010). The future of mobile TV: when mobile TV meets the internet and social networking. *In Mobile TV: Customizing content and experience* (pp. 305-326). Springer London.

Montpetit, M. J., Klym, N., & Mirlacher, T. (2011). The future of IPTV. *Multimedia Tools and Applications*, 53(3), 519-532.

Yarali, A., & Cherry, A. (2005, November). Internet protocol television (IPTV). *In TENCON 2005 2005 IEEE Region 10* (pp. 1-6). IEEE.

Lee, S. B., Muntean, G. M., & Smeaton, A. F. (2009). Performance-aware replication of distributed pre-recorded IPTV content. *IEEE Transactions on Broadcasting*, 55(2), 516-526.

Li, M., & Wu, C. H. (2010). *A cost-effective resource allocation and management scheme for content networks supporting IPTV services*. Computer Communications, 33(1), 83-91.

Dukes, J., & Jones, J. (2004, November). Using dynamic replication to manage service availability in a multimedia server cluster. *In International Workshop on Multimedia Interactive Protocols and Systems* (pp. 194-205). Springer Berlin Heidelberg.

Sujatha, D. N., Girish, K., Venugopal, K. R., & Patnaik, L. M. (2008, January). An efficient storage mechanism to distribute disk load in a VoD server. *In International Conference on Distributed Computing and Networking* (pp. 478-483). Springer Berlin Heidelberg.

Lie, P. W., Lui, J. C., & Golubchik, L. (2000). Threshold-based dynamic replication in large-scale video-on-demand systems. *In Continuous Media Databases* (pp. 31-58). Springer US.

Little, T. D., & Venkatesh, D. (1993, November). Probabilistic assignment of movies to storage devices in a video-on-demand system. *In International Workshop on Network and Operating System Support for Digital Audio and Video* (pp. 204-215). Springer Berlin Heidelberg.

Sobe, A., Elmenreich, W., & Böszörmenyi, L. (2010, October). Towards a self-organizing replication model for non-sequential media access. In Proceedings of the 2010 ACM workshop on Social, *adaptive and personalized multimedia interaction and access* (pp. 3-8). ACM.

Huang, Y. F., & Fang, C. C. (2004). Load balancing for clusters of VOD servers. *Information Sciences*, 164(1), 113-138.

Zhou, X., & Xu, C. Z. (2002, April). Request Redirection and Data Layout for Network Traffic Balancing in Cluster-Based Video-on-Demand Servers. In ipdps.

Bikfalvi, A., García-Reinoso, J., Vidal, I., Valera, F., & Azcorra, A. (2011). P2P vs. IP multicast: Comparing approaches to IPTV streaming based on TV channel popularity. *Computer Networks*, 55(6), 1310-1325.

Joe, I., Yi, J. H., & Sohn, K. S. (2011). A content-based caching algorithm for streaming media cache servers in cdn. *In Multimedia, Computer Graphics and Broadcasting* (pp. 28-36). Springer Berlin Heidelberg.

García, R., Pañeda, X. G., Melendi, D., & Garcia, V. (2009). Probabilistic analysis and interdependence discovery in the user interactions of a video news on demand service. *Computer Networks*, 53(12), 2038-2049.

Cho, D. H., Lee, K. Y., Choi, S. L., Chung, Y. D., Kim, M. H., & Lee, Y. J. (2008). A request distribution method for clustered VOD servers considering buffer sharing effects. *Journal of Systems Architecture*, 54(1), 335-346.

Meng, S., Liu, L., & Yin, J. (2013). A collaborative and scalable platform for on-demand IPTV services. *IEEE Transactions on Services Computing*, 6(3), 358-372.

Kanrar, S. (2012). *Analysis and implementation of the large scale video-on-demand system*. arXiv preprint arXiv:1202.5094.

Wong, E., M. (2004). *Method for Estimating the Number of Concurrent Users*, report. Retrieved from: http://wenku.baidu.com/view/ 68f1853a5802-16fc700afd74.html.

Nafaa, A., Murphy, S., & Murphy, L. (2008). Analysis of a large-scale VOD architecture for broadband operators: a P2P-based solution. *IEEE Communications Magazine*, 46(12).

Jain, Y. K., & Bhandare, S. K. (2011). Min max normalization based data perturbation method for privacy protection. *International Journal of Computer & Communication Technology (IJCCT)*, 2(8), 45-50.

Gaber, S. M. A., & Sumari, P. (2014). Predictive and content-aware load balancing algorithm for peer-service area based IPTV networks. *Multimedia tools and applications*, 70(3), 1987-2010.

Huang, C., Zhou, G., Abdelzaher, T. F., Son, S. H., & Stankovic, J. A. (2005, December). Load balancing in bounded-latency content distribution. In Real-Time Systems Symposium, 2005. RTSS 2005. *26th IEEE International* (pp. 12-pp). IEEE.

Fati, S., Sumari, P., & Budiartu, R. (2015). Fair and popularity based content allocation scheme for IPTV delivery networks. Int J Comput Appl, 131(5), 21-26.

Fati, S. M., Budiartu, R., & Sumari, P. (2014, January). Provisioning virtual IPTV delivery networks using hybrid genetic algorithm. *In Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication* (p. 106). ACM.

Fati, S. M., & Sumari, P. (2016). Content-aware replica placement strategy for IPTV services over peer-service area architecture. *Multimedia Tools and Applications,* 1-25.