

Resource Usage Prediction Models for Optimal Multimedia Content Provision

Yiannos Kryftis, *Member, IEEE*, George Mastorakis, *Member, IEEE*,
 Constandinos X. Mavromoustakis, *Member, IEEE*, Jordi Mongay Batalla, *Member, IEEE*,
 Joel J. P. C. Rodrigues, *Senior Member, IEEE*, and Ciprian Dobre, *Member, IEEE*

Abstract—This paper proposes a network architecture that utilizes novel resource prediction models for optimal selection of multimedia content provision methods. The proposed research approach is based on a prototype system, which exploits a resource prediction engine (RPE), utilizing time series and epidemic spread models, for optimal and balanced distribution of the streaming data among content delivery networks, cloud-based providers and home media gateways. The proposed epidemic diseases models adopt the characteristics of the multimedia content delivery over the network architecture. In this context, this paper aims to present the advantages of using such models, by presenting and analyzing an epidemic spread scheme for video-on-demand (VoD) delivery, to predict future epidemic spread behavior. In addition, this paper presents two algorithms, adopted in the prototype network architecture, for optimal selection of multimedia content delivery methods, as well as balanced delivery load, by exploiting the RPE. Both algorithms and models are evaluated to establish their efficiency, toward effectively predicting future network traffic demands. The simulation results verify the validity of the proposed approach, identifying fields for future research and experimentation.

Index Terms—Content delivery networks (CDNs), epidemic spread models, media distribution middleware (MDM), multimedia services systems, network architectures, quality of experience (QoE), resource prediction engine (RPE).

I. INTRODUCTION

GIVEN the tremendous evolution of multimedia-related technologies over the Internet, more pressure is applied

Manuscript received January 19, 2015; revised November 27, 2015; accepted February 20, 2016. This work was supported in part by the European Union, Eurostars Programme, under Project 8111, in part by DELTA “Network-Aware Delivery Clouds for User Centric Media Events,” in part by the Instituto de Telecomunicações, in part by the Next Generation Networks and Applications Group (NetGNA), Portugal, and in part by the National Funding from the FCT—Fundação para a Ciência e a Tecnologia through the UID/EEA/50008/2013 Project.

Y. Kryftis and C. X. Mavromoustakis are with the Department of Computer Science, University of Nicosia, 1700 Nicosia, Cyprus (e-mail: kryftis.y@unic.ac.cy; mavromoustakis.c@unic.ac.cy).

G. Mastorakis is with the Department of Informatics Engineering, Technological Educational Institute of Crete, 71500 Heraklion, Greece (e-mail: gmastorakis@staff.teicrete.gr).

J. M. Batalla is with the National Institute of Telecommunications, Warsaw University of Technology, 04-894 Warsaw, Poland (e-mail: jordim@interfree.it).

J. J. P. C. Rodrigues is with the Instituto de Telecomunicações, University of Beira Interior, 6201-001 Covilhã, Portugal, and also with the University of Salamanca, 37008 Salamanca, Spain (e-mail: joeljr@ieee.org).

C. Dobre is with the Faculty of Autonomic Control and Computers, University Politehnica of Bucharest, 060042 Bucharest, Romania (e-mail: ciprian.dobre@cs.pub.ro).

Digital Object Identifier 10.1109/JSYST.2016.2548423

for further research and development on the field of multimedia content distribution. A significant part of the global Internet traffic is generated by video- and audio-on-demand services or other multimedia services, while the amount of this traffic is expected to double in the future [1], leading toward the Future Media Internet. Recent advances in connected media technologies and social networks are the driving forces, while broadband infrastructure growth and cloud computing that has emerged as a new paradigm for hosting and delivering services over the Internet are the keystones for the upcoming Future Media Internet. The rapidly transforming environment that surrounds the citizens forces them to the demand of more community-centric experiences through networked/connected media and social networks and to even better quality of experience (QoE). Media content delivery plays a key role for the QoE, pressing for more research on novel network architectures, as well as the relative components that allow efficient and balanced content delivery. In addition, novel algorithms and models for the prediction of the resources are vital to be adopted for efficient multimedia content provision.

Tackling such challenges, this paper goes beyond the current state-of-the-art, elaborating on a new multimedia services delivery solution. The proposed solution is based on the optimum allocation of the resources used for content transmission to efficiently satisfy different users’ requests through the exploitation of existing servers’ infrastructures capabilities. Such capabilities are available in conventional clouds (i.e., public or private computing infrastructure configurations, usually offered by over-the-Top providers) and in content delivery networks (CDNs). Additionally, they can be offered by home media gateway clouds [i.e., home gateways/community gateways configurations, exploited in peer-to-peer (P2P) mode]. The proposed approach foresees a new business model in multimedia services delivery over the Internet, strongly but smoothly leveraging (in an evolutionary way) new mechanisms and systems. Among others an epidemic spread model is proposed that is proven to accurately describe a video-on-demand (VoD) spread.

In this context, this paper is organized as follows. Section II presents related work on resource prediction engines (RPEs) and on epidemic models, as well as the research motivation of this paper. Section III elaborates on the proposed research approach based on a novel network architecture that utilizes an RPE, an epidemic model, and two algorithms for optimal multimedia services provision. Finally, Section IV provides the

evaluation results and Section V includes the conclusion of this paper, accompanied with fields for future research.

II. RELATED WORK AND RESEARCH MOTIVATION

Several existing research attempts elaborate on the combination of different delivery methods, in order to achieve better QoE for the users. In [2], Xu *et al.* propose a CDN-P2P hybrid architecture for cost-effective streaming media distribution that combines the advantages of using CDN for providing high QoE with the low cost of using P2P-based stream. Yin *et al.* [3] present the design and deployment of a hybrid CDN-P2P system for live video streaming, demonstrating the improvement in startup delay time and in stability. In [4], Ciullo *et al.* also propose a peer-assisted video distribution in order to reduce the server workload and to introduce scalability to the system. They suggest a stochastic fluid framework that allows the estimation of the needed bandwidth for the satisfaction of the user requests based on predefined scenarios. In an energy-aware approach in [5], Mandal *et al.* analyzed and presented the advantages of the integration between CDN and P2P networks [6]–[9]. Zhang *et al.* conducted a measurement study on Kankan, which is one of the leading VoD streaming service providers in China and is based on a hybrid CDN-P2P architecture. They present how the provider utilizes the P2P network for storage and streaming of videos, and how the CDN servers assist the streaming procedure [10]. Current research approaches focus on how to benefit from the combination of the different delivery methods but they do not take consideration of handling each resource separately. In comparison to such approaches, our proposed solution goes beyond the current state-of-the-art, by handling each resource (i.e., streaming channel) separately based on the prediction of the future demand for each resource, as well as the predicted network metrics. The early prediction before the actual need provides to the proposed system the ability to enforce management actions to maintain a high QoE for the users [11], [12].

An RPE constitutes an important part of multimedia content delivery system, in order to offer the desired QoE to the end users [13]–[15]. Its role is to provide the ability to efficiently predict the needed bandwidth capacity and the upcoming network fluctuations. The prediction engine has to be based on novel methods and models that can accurately forecast the future demands, in order to trigger through a management plane the proper actions for keeping the desired quality for the streaming sessions. In [16], Niu *et al.* present some time-series analysis techniques to predict the server bandwidth demand and the peer upload for content delivery in peer-assisted VoD services. The analysis includes the prediction of future population for each video channel, by analyzing and fitting to existing models, past data about the population of video channels. The seasonal autoregressive integrated moving average (ARIMA) model [17] is exploited, for avoiding the periodicity. Additionally, they use machine learning techniques for inferring the initial population of a newly released channel, by utilizing pass data from newly released videos as training data. For the prediction of the server bandwidth demands by a video channel at future time, the autoregressive moving-average (ARMA) model was used [17]. They prove that the entire procedure has reasonable

computation cost. In [18], Niu *et al.* present a system for VoD providers in the Cloud that provides the ability to predict the upcoming need for bandwidth in order to autoscale accordingly. The near future demands' expectations are estimated based on the history of demands as monitored by the cloud monitoring services. This provides the opportunity to reserve the minimum bandwidth needed for satisfying the demand in the desired quality. In a similar manner, our proposed RPE utilizes a combination of statistical models for the prediction of future needs. The innovation originates from the selection of the appropriate model, out from a pool of statistical models (ARMA [17], ARIMA [17], theta method [19], and cubic splines [20]), which better describes each content delivery. Additionally, the epidemic model proposed by this paper provides to the RPE the capability to predict sudden and intense increase of delivery need for specific content. For an RPE able to forecast future demands, the recent advances in connected media technologies and social networks should be taken into account. Social networks play a significant role in content delivery, by providing ways of interactions among users that can lead to a lightning spread of content [21]–[23]. In [24], Gonçalves *et al.* suggested a probabilistic resource provisioning approach that utilizes the basic susceptible-infectious-recovered (SIR) model, developed for epidemiology spreading, to represent the sudden and intense workload overflow in VoD delivery process. More specifically, they use Markov chain to describe the behavior of the users and they trace the cases of epidemic spread, or as the call it buzz effect, by introducing a hidden Markov model with two different rates to represent the buzz and buzz-free behavior. Contrary, we propose an epidemic model with more states, customized to describe the epidemic spread of content through the proposed overall system. Additionally, our suggestion for the prediction is based on fitting historical data into the proposed model. Epidemic model spreading in scale-free networks has been intensively studied [25]. In [26], Pastor-Satorras and Vespignani analyzed data from computer virus infections and they defined a dynamical model for the spreading of infections on scale-free networks. The spread of computer viruses really resembles the epidemic spread of human diseases. In this context, this paper extends the basic disease models, by presenting a novel model that can be used to accurately describe the multimedia content delivery (i.e., VoD delivery) and so it can help forecasting epidemic spread of multimedia content. The majority of disease models are based on a splitting in compartments of the individuals in a population based on their disease status [25], [27]–[29]. The basic SIR model provides the foundations of almost all mathematical epidemiology. The differential equations that describe the model are as follows:

$$\frac{dS}{dt} = b \cdot N - \beta \cdot S \cdot I - d \cdot S \quad (1a)$$

$$\frac{dI}{dt} = \beta \cdot S \cdot I - \delta \cdot I - d \cdot I \quad (1b)$$

$$\frac{dR}{dt} = \delta \cdot I - d \cdot R \quad (1c)$$

$$S + I + R = N. \quad (1d)$$

In these equations, S , I , and R refer to the number of susceptible, infectious, and recovered individuals, respectively, in a

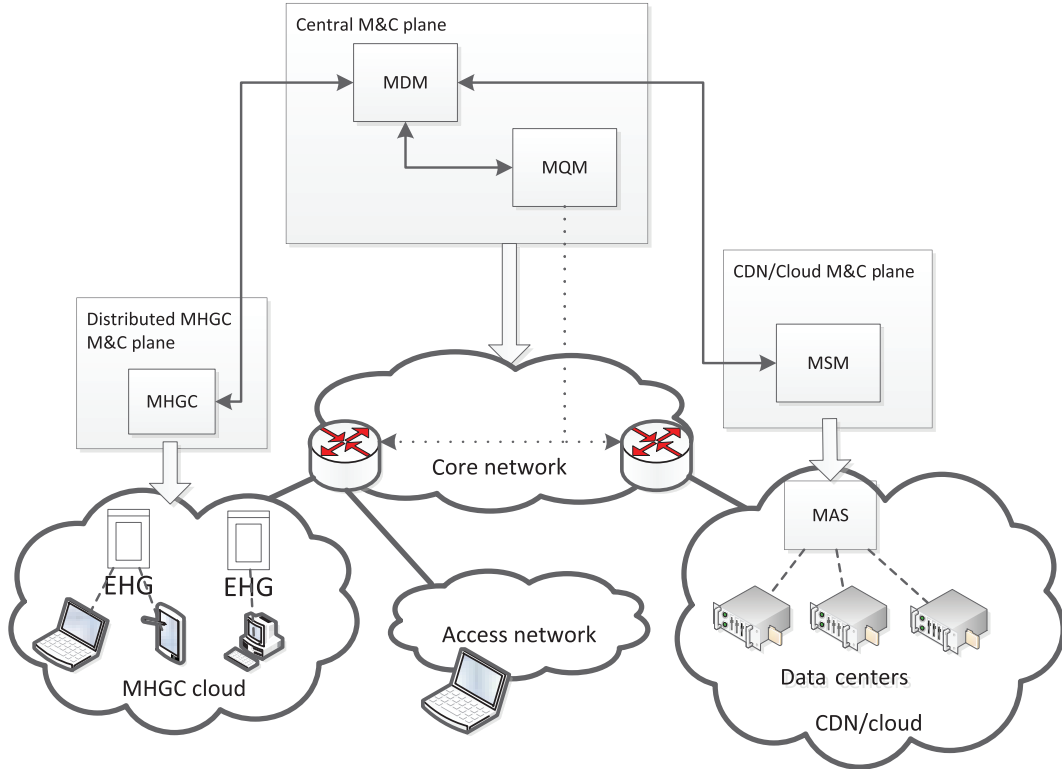


Fig. 1. Proposed network architecture.

population of size N . The other parameters are the birth rate b , the natural death rate d , and the rate of recovery from infection δ . The force of infection π , is the rate, at which susceptible individuals become infected. It is a function of the number of infectious individuals; this parameter contains information about the interactions between individuals that lead to the transmission of infection. When the population is randomly mixing, the force of infection can be calculated as follows:

$$\pi = \beta \cdot \frac{I}{N} \quad (1e)$$

where β is the effective number of contacts per unit time. This leads to a nonlinear term ($\frac{\beta \cdot S \cdot I}{N}$) representing the transmission of infection, generating a variety of rich dynamical behaviors. Theoretical modeling of how diseases spread in complex networks is based on the assumption that the propagation is driven by reaction processes and that the transmission occurs from every infected neighboring entity at each time step, producing a diffusion of the epidemics on the network. Possible modifications on the available states of the SIR model lead to some widely used epidemic models like the susceptible-infectious-susceptible (SIS) model where the recovered state does not exist and individuals are considered immediately susceptible. In the same context, the maternally derived immunity-SIR (MSIR) model compared to the SIR model includes a state for a population born with immune to the disease. In such a model, an additional differential equation is needed to describe the transitions from state M to the state S. This equation takes into consideration the percentage of population with immunity and its lasting period [30], [31]. Similar models have been

used in finite-size scale-free networks, for traffic-driven epidemic spread [32], for efficient data streaming [33] and for virus spread in such networks [34]. Also, this approach has been used in the area of multimedia content distribution [21] in dynamic resource management [24], and for segmented file sharing [35]. Although there is impressive research in epidemic models for use for the description of behavior of computer networks, there is a lack of a model, able to describe the specific need for multimedia content delivery architectures. Part of this paper is dedicated to an epidemic model for that purpose. Additionally, this paper elaborates on a network architecture that predicts the future content delivery demands and the future network usage, by utilizing novel models and algorithms, performing all the necessary adaptations to deliver the content in an optimal and balanced way for the optimal provision of the desired quality of service (QoS) and QoE to the end users.

III. EPIDEMIC MODEL FOR OPTIMAL MULTIMEDIA SERVICES PROVISION USING AN RPE

A. Network Architecture

The introduction of an RPE with epidemic models utilization, toward the provision of multimedia services, demands a network architecture with management components in different layers that cooperate during the multimedia delivery process, achieving optimal content delivery. The proposed network architecture is shown in Fig. 1. The upper layer is called central management and control (M&C) plane and it coordinates the collaboration environment, by interchanging information with existing M&C planes of the CDN and cloud providers, as

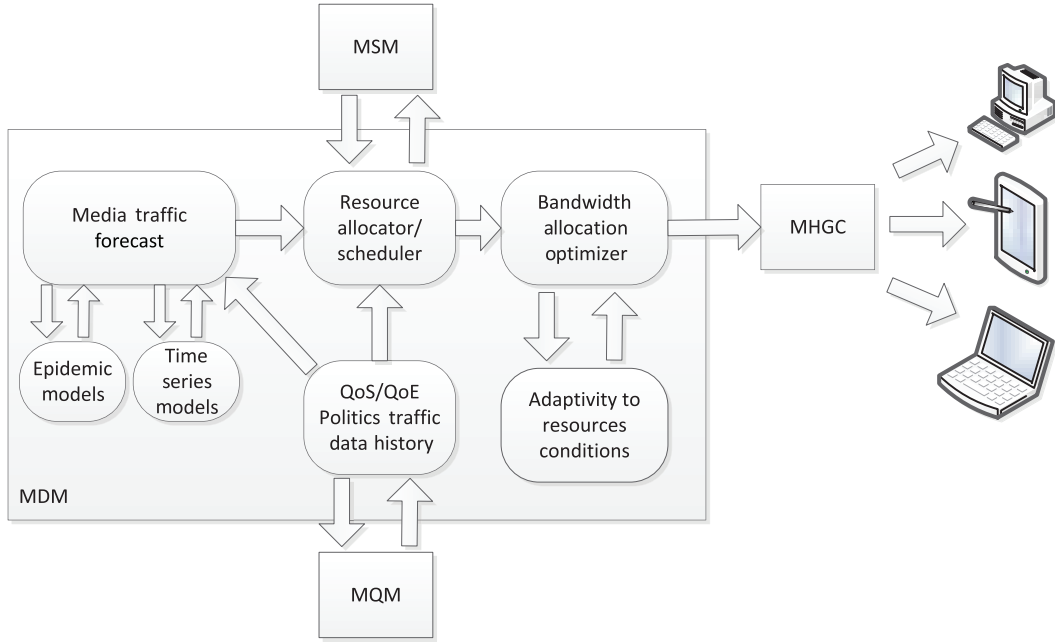


Fig. 2. MDM internal architecture.

well as the distributed M&C plane of the media home gateway cloud (MHGC) provider, consisted of user gateways, forming a P2P network. The proposed network architecture as presented in Fig. 1 consists of the following entities: a media distribution middleware (MDM), a media QoE meter (MQM), a media services manager (MSM), an enhanced home gateway (EHG), and a media advanced streamer (MAS). In a bottom to top presentation of components based on the functionality, the EHG entity is part of the home equipment of the end users. The control modules of EHG's constitute the MHGC M&C plane and they are responsible for creating the MHGC ad hoc system from a set of P2P connected EHG's. Each EHG receives content requests from the users, requesting data from the MDM, about which MHGC peers should get involved to efficiently deliver requested content. EHG collaborates with MSM entities that reside in CDN/cloud M&C planes and manage all service providers resources, to obtain media content requested by the user, if the content is not stored on any of EHG's belonging to given MHGC. The MSM, according to the recommendations received from the MDM, takes a decision on which server should stream the requested media and with which bitrate. In this way, the MSM, contrary to the existing solutions, performs adaptation decision, taking into account not only the available bandwidth, but also considering other important information addressed by the MDM, such as the estimated QoE value and the prediction of the potential upcoming streaming sessions. The MAS entity resides in the CDN/cloud domain as a standalone component. Its role is to perform the streaming process, according to the instructions received from the MSM/EHG entity. MQM component is responsible for continuous monitoring of network metrics at the users and the service providers domain access points, as well as the users context and preferences. Based on the data gathered by the set of the MQM probes, distributed all over the domain, this entity provides to the MDM the related data about the current

network conditions and the estimated value of QoE available for a user. Moreover, the MQM sends alerts to the MDM, only if any of the monitored QoS/QoE parameters declines below the allowed level. MDM is the main component of the central M&C plane. It executes all necessary operations and determines all data required for optimal allocation of the available resources at each resource providers domain. As a result, the MDM returns guidelines, which resources should be used for handling given users request, to achieve the best (in terms of efficiency) resource exploitation.

B. RPE for Optimal Multimedia Services Provision

The MDM adopts an RPE, in order to be able to predict future demands for resources. The prediction is divided into long-term prediction for future demands for resources and short-term prediction for some important network metrics like throughput. The long-term prediction takes as input the demand for each resource in the past, using a combination of statistical methods and algorithms for the adaptivity to description models, in order to predict future demands. This provides the opportunity to the system to make the optimal distribution of data in clouds, CDNs, and EHG's based on the prediction before the actual need. On the other hand, the short-term prediction is used for predicting and preventing upcoming network congestion issues, by triggering the proper management actions. Fig. 2 presents the internal architecture of the MDM component. The QoS/QoE politics traffic data history component collects and stores the monitoring data delivered from the MQM. It forwards it periodically to the Media Traffic Forecast to generate the prediction for the traffic in the network. The Media Traffic Forecast utilizes the epidemic model for the prediction of the upcoming epidemic or not spread of the content and the time-series models for the prediction of future values of the resources. The

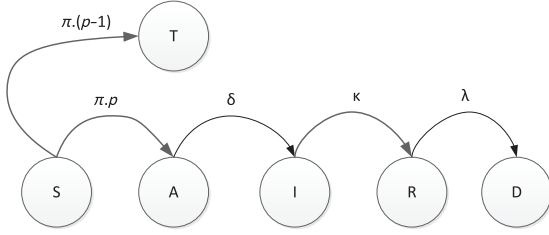


Fig. 3. Epidemic model for efficient multimedia content provision.

outcome of the forecasts is used as input to the resource allocator/scheduler. It uses algorithms that combine the current and predicted values of specific metrics to decide on the optimal delivery methods and the most suitable servers to perform the multimedia content delivery. The results are feeding the MSM component, as recommendations regarding which server should stream the requested media, while at the same time the bandwidth allocation optimizer calculates the optimal bandwidth allocation for the P2P delivery between the MHGC devices. It is an online system, which takes into consideration the network metrics that come from the MQM to decide on the most suitable peers to perform the streaming. It delivers that information directly to the MHGC devices. MHGC devices exchange management information between them and together they constitute an M&C plane that manages the P2P network between the EHG devices. The adaptivity to resources conditions component ensures that the bandwidth allocation optimizer runs periodically to retain the optimal suggestions based on updated information from the resources and network.

C. Epidemic Models for Prediction of VoD Download Rate

This paper examines the effectiveness of the epidemic models on the prediction of VoD usage as part of the general issue of optimal content delivery. As shown in Fig. 3, the model divides the population into several compartments based on the percentage of the population in each state. The susceptible (S) group includes subscribers that can download the video, the active (A) includes the users that are currently downloading the video, the infected (I) contains the users that downloaded the video and they can spread it through social networks (if they liked the video), the recovered (R) contains the users that passed from the infected phase but they do not spread the video any more (after some period of time) while the deleted (D) group includes users that removed the video after sometime or the video was automatically removed from the cache after some period of time. The turned_down (T) group includes the users that belonged to the S category and they took the decision to turn down the Video, so they will never download it. In the model, $N_S(t)$, $N_A(t)$, $N_I(t)$, $N_R(t)$, $N_D(t)$, $N_T(t)$, $t \geq 0$ are stochastic processes representing the time evolution of each population.

Suppose there are n clients of the VoD provider, they all belong to the group S at time $= 0$, when a VoD is initially uploaded by the provider. The transition rate from state S to A consists of the probability to have a new spontaneous viewer, plus the probability to have some users that learned about the

video from their social contacts and they came to a decision to watch the video. So, (1e) of the SIR model that expresses the nonlinear term of the rate of transmission of infection can be extended to

$$\pi(t) = \gamma + \beta \cdot N_S(t) \cdot N_I(t) \quad (2a)$$

where β is the social network contact rate for users based on the specific video, and γ is the number of spontaneous viewers that in some cases can be considerable important since a specific VoD may be advertised and promoted by the VoD provider. For the transition from active state to infected state, there is a need of consideration of the download rate of users and of length of the video. Since the video download is happening directly during viewing and there are mechanisms for balanced delivery, the time needed for the transition is considered as a random variable and is expressed as a Poisson process with mean value of the duration of the video. The transitions, from infected to recovered and from recovered to deleted, are also considered as random variables since each user can spread the information to its social network for a random period of time and can also keep the video in its EHG device again for a random period of time. So, for the proposed model, the specific transitions are again described as Poisson processes with mean times the estimations of how long the users are spreading the information to their social networks and the period that the video stays in the device of each user. The rates can be expressed as follows:

$$\delta(t) = \frac{1}{\text{video duration}} \quad (2b)$$

$$\kappa(t) = \frac{1}{\text{spread period}} \quad (2c)$$

$$\lambda(t) = \frac{1}{\text{keep in cache period}} \quad (2d)$$

This model makes the following assumptions. 1) A user that downloads a video will never request it again. 2) There are no changes in user's population/subscribers. 3) Users in T state that turned down the video will never become susceptible again. The first two assumptions help on making the analysis simpler without losing the generality, since they are well fitting a VoD provider use case. Regarding the first assumption, when a user downloads a video, it remains in its local EHG for a few days. So, a redownload is performed only in the case that a user wishes to view it again after that period. This case does not importantly affect the analysis, since the epidemic spread occurs in short periods of time and with the existence of cache it is unlikely to have double downloads during the study of a single epidemic spread incidence. Similarly, the second assumption is realistic since the population of the subscribers remains almost constant within such short periods. The third assumption does not really change the analysis, since it is not possible to calculate the population of state T before the actual spread of the video. In the measurements, S and T populations are handled together and they can be separated only in the end of each VoD life cycle. The following equations describe the model:

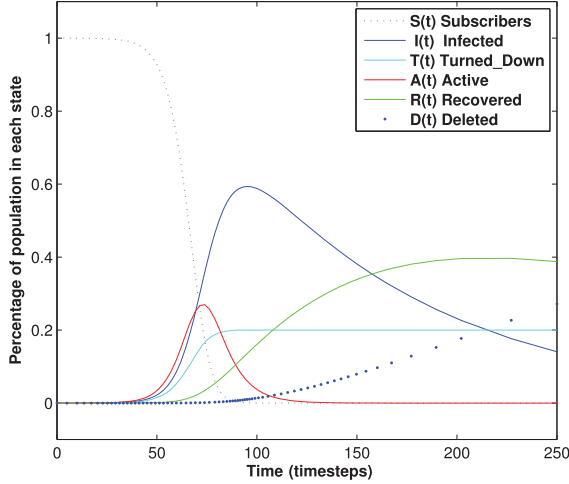


Fig. 4. Percentage of population in each state.

$$\frac{dS}{dt} = -(\beta \cdot I + \gamma) \cdot S \quad (3a)$$

$$\frac{dT}{dt} = (\beta \cdot I + \gamma) \cdot S \cdot (1 - p) \quad (3b)$$

$$\frac{dA}{dt} = (\beta \cdot I + \gamma) \cdot S \cdot p - \delta \cdot A \quad (3c)$$

$$\frac{dI}{dt} = \delta \cdot A - \kappa \cdot I \quad (3d)$$

$$\frac{dR}{dt} = \kappa \cdot I - \lambda \cdot R \quad (3e)$$

$$\frac{dD}{dt} = \lambda \cdot R \quad (3f)$$

$$S + T + A + I + R + D = 1. \quad (3g)$$

Fig. 4 presents simulation results executed in MATLAB. The aforementioned differential equations were expressed as MATLAB equations and the simulation was performed in timesteps. After each timestep, part of the population was moving to a different state based on the equations and the rates. In simulation results of Fig. 4, the rates are as follows: $\beta = 0.5$, $\delta = 0.1$, $\kappa = 0.01$, $\lambda = 0.005$, $p = 0.8$. The whole population belongs to the S state ($S = 1$) when $t = 0$. At time = 50 timesteps, the population of active (A) increases showing that there are active users downloading the video, while at time = 70 timesteps the population of Infected (I) increases significantly. An important outcome is the final population in state turn down (T) that tends to $p - 1 = 0.2$, while the population of deleted (D) tends to $p = 0.8$ as expected since 80% of the population chooses to view the video. The most important line is the $A(t)$ since it depicts the bandwidth need for covering the needs of the active downloads. It is clear that in case of epidemic spread of a specific video, the population of simultaneous downloads is significantly increased, something that increases the difficulty in delivering high QoS. A solution to the problem is the use of P2P delivery complementary to the cloud and CDN delivery. An important observation is that by the time when the active (A) users introduce a significant increase in their population, there is always an important number of users in infected (I) and recovered (R) states that can seed the video for the them

through P2P delivery method. Finally, it is clear that the transmission from recovered (R) state to deleted (D) does not affect the needed recovered population when it is most needed.

Lemma 1: if $\gamma > 0$ then

$$\lim_{t \rightarrow \infty} (T(t) + D(t)) = 1$$

Proof: $T(t) = T(t) - T(0) = \int_0^t \frac{dT}{dt} dt$. Since $\frac{dT}{dt} = -(1-p) \frac{dS}{dt}$, we conclude that $T(t) = -\int_0^t (1-p) \frac{dS}{dt} dt = -(1-p)(S(t) - S(0)) = (1-p)(1 - S(t))$. Notice that $\frac{dD}{dt} \geq 0$, hence D is increasing. $D(s) \leq 1$ for all s therefore $\lim_{s \rightarrow \infty} \frac{dD}{dt}|_s = 0$. Hence, $\lim_{s \rightarrow \infty} R(s) = 0$ if we assume $\lambda > 0$. Similarly,

- 1) $\lim_{s \rightarrow \infty} \frac{dR}{dt}|_s = 0$, and therefore $\lim_{s \rightarrow \infty} I(s) = 0$ (assuming that $\kappa > 0$);
- 2) $\lim_{s \rightarrow \infty} \frac{dI}{dt}|_s = 0$, and therefore $\lim_{s \rightarrow \infty} A(s) = 0$ (assuming that $\delta > 0$);
- 3) $\lim_{s \rightarrow \infty} \frac{dA}{dt}|_s = 0$, and therefore $\lim_{s \rightarrow \infty} S(s) = 0$ (assuming that $\gamma > 0$).

Hence, since $S + T + A + I + R + D = 1$, we see that $\lim_{s \rightarrow \infty} T(s) + D(s) = 1$. By the equation $T(t) = (1-p)(1 - S(t))$, we see that $\lim_{s \rightarrow \infty} T(s) = 1 - p$. Hence, we also see that $\lim_{s \rightarrow \infty} D(s) = p$.

The outcome of Lemma 1 is the fact that in infinite time all the population will finally go to the state T-turn down and D-deleted, given we have $\gamma > 0$. The lemma is useful for the analysis after the end of the content delivery for extracting the percentage of users that turned down the video.

An important metric for the model is the basic reproduction rate of the epidemic R_0 that can be calculated as in (global analysis of multistrains SIS, SIR, and MSIR epidemic models)

$$R_0 = \frac{p * \beta}{\kappa}.$$

Lemma 2: if $R_0 > 1$ then the epidemic cannot maintain itself.

Proof: If $R_0 > 1$, then on average, each infected individual infects more than one other member of the population and a self-sustaining group of infectious individuals will propagate. If $R_0 < 1$, then the epidemic cannot maintain itself because each individual, on average, infects less than one member of the population. ■

The estimation of R_0 is not easy before the actual spread of the VoD because the β value is not easy to be predicted, since the social impact of each VoD is different. An estimation can be done based on some of the video properties (category, actors, director, etc.) and some history data, but this is out of the scope of the specific paper.

It is interesting to study and compare the model in the case where there is no epidemic spread ($R_0 < 1$). To simplify the analysis, it is safe to say that if there is no epidemic spread it can be considered that $\beta = 0$. In this case, the infected (I) and recovered (R) population are in the same state captured (C). The equations are transformed as follows:

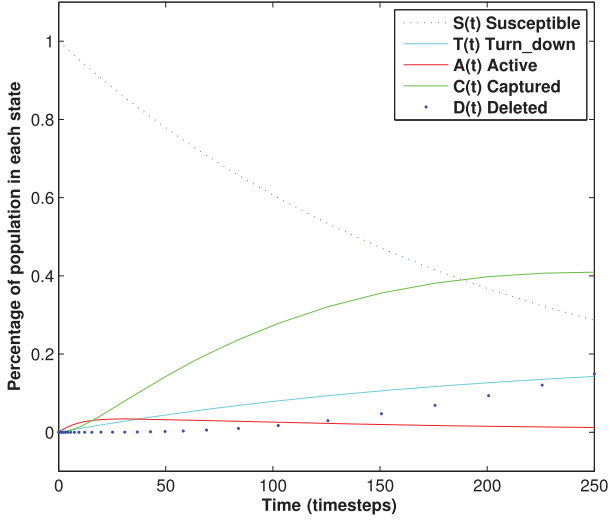


Fig. 5. Simulation results without epidemic spread.

$$\frac{dS}{dt} = -\gamma \cdot S \quad (4a)$$

$$\frac{dT}{dt} = \gamma \cdot S \cdot (1 - p) \quad (4b)$$

$$\frac{dA}{dt} = \gamma \cdot S \cdot p - \delta \cdot A \quad (4c)$$

$$\frac{dC}{dt} = \delta \cdot A - \lambda \cdot C \quad (4d)$$

$$\frac{dD}{dt} = \lambda \cdot C \quad (4e)$$

$$S + T + A + C + D = 1. \quad (4f)$$

By solving the first-order differential equations for S and A with consideration that $S(0) = 1$ and $A(0) = 0$ the outcome is

$$S(t) = e^{-\gamma \cdot t} \quad (5a)$$

$$A(t) = \frac{\gamma \cdot p}{\delta - \gamma \cdot p} (e^{-\gamma \cdot t} + e^{-\delta \cdot t}). \quad (5b)$$

Fig. 5 presents simulation results when the rates are as follows: $\delta = 0.1$, $\kappa = 0.01$, $\lambda = 0.005$, $p = 0.8$. The difference it is obvious since the active (A) population remains very low through the whole delivery process. If we consider that the system is able to deliver simultaneously to up to a specific percentage of users, then the value of γ can be modified. The VoD provider is able to manage the advertise, or the position in the menu of each VoD and so it can modify the γ value of the spontaneous users.

Lemma 3: Prediction based on the model. If the $A(t)$ is exponential then it is epidemic, if the $A(t)$ is polynomial then it is not epidemic. In case of predicting an epidemic spread of the video the algorithm is modified.

Proof: In model with epidemic spread

$$\frac{dA}{dt} = (\beta \cdot I + \gamma) \cdot S \cdot p - \delta \cdot A \approx \beta \cdot I \cdot S \cdot p.$$

In model without epidemic spread

$$\frac{dA}{dt} = \gamma \cdot S \cdot p - \delta \cdot A \approx \gamma \cdot S \cdot p. \quad \blacksquare$$

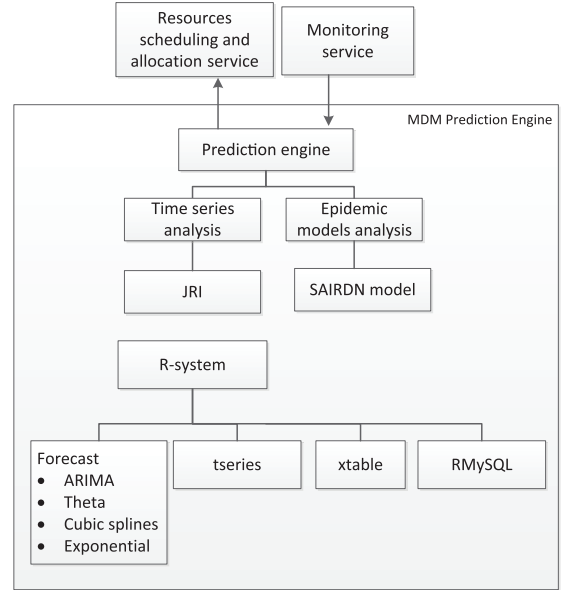


Fig. 6. Implementation architecture of the RPE.

The multiplication of S and I populations is what causes the epidemic spread. An early perception of such behavior could benefit the content delivery system.

D. Resource Prediction Algorithms With Use of Epidemic Models

An internal view of the implementation architecture of the prediction engine as part of the Media Traffic Forecast is presented in Fig. 6. Input comes from the monitoring service and more specifically the MQM through the QoS/QoE politics traffic data history component. The prediction engine is implemented in Java and it is divided into two parts based on the functionality concerning time-series models and epidemic models. For the implementation of time-series models, there is a use of the JRI, Java Interface [36] for the interactions with the R-system [37]. R is a very popular free software environment for statistical computing and graphics. The standard stats package of R-system includes multiple time-series models and prediction methods. The forecast package [38], [39] of R implements automatic forecasting with multiple methods, including ARIMA models, exponential smoothing methods, Theta method [19], cubic splines [20], and many others. Hyndman and Khandakar in [39] present the implementation of exponential smoothing methods and the ARIMA modeling approach in the forecast package. The proposed prediction engine uses the aforementioned packages, extending them in order to achieve optimal prediction. If a long-term (i.e., for the next days) prediction has to be achieved for the needed bandwidth of a specific VoD, the prediction engine needs to exploit the history of the bandwidth reservation for the specific VoD. The history data are utilized for fitting in the proper statistical model, suitable for the specific VoD and then the forecast function performs the prediction. The result is the estimation of the need of bandwidth for the specific VoD after 1 h. This value feeds the resource allocator/scheduler to decide how to serve

the estimated future need. The epidemic model is implemented in Java. It examines the history data behavior to conclude if it resembles the exponential function something that would predict that the VoD spreads with epidemic speed based on Lemma 3. The MDM component uses the predicted future values for the metrics, in order to take the decisions for delivery of requested media, which may be streamed: 1) directly from the cloud; 2) through deployed surrogate servers of the CDN; 3) by establishing an MHGC *ad hoc* system and using a combined P2P-based technology of distribution with multisource, multi-destination congestion control algorithms; or 4) a combination of parts or all of them (thanks to stream-switching adaptation technique). The results are forwarded to the MSM component that is responsible for the actual streaming of the data to the user. The selection algorithm is presented. Algorithm 1 gets as input the current and predicted bandwidth needed for the delivery of specific content over the system and triggers the proper mechanisms to accomplish it in an optimal way. The **neededBW** variable that represents the expected needed bandwidth for the overall delivery of a specific VoD, takes the higher value among the current and the predicted bandwidth need as calculated by the prediction engine. The **epidemicSpread** boolean variable is assigned a TRUE or FALSE value based on the behavior similarity to the exponential function that indicates an epidemic spread (Lemma 3). If *epidemicSpread* = TRUE then the algorithm uses all delivery methods simultaneously and it informs the VoD provider to reduce the advertisement of the specific VoD. If *epidemicSpread* = FALSE it makes the selection based on preset thresholds for bandwidth usage, based on administrative high level decisions and network status. If **neededBW** is below the low threshold, only the cloud delivery will be used. If it is above, CDN servers will be used. The number of CDNs to be used is defined as a percentage of the available servers, based on the distance of **neededBW** from the low and high threshold. After the high threshold is reached, a P2P delivery method is exploited. This algorithm provides the advantage that the data are not distributed before the actual need. In the case of a VoD with low customers demand, the CDNs will not be used for its distribution. On the other hand, if a VoD becomes viral and it is spreading epidemically, an early prediction will occur that will allow to use all the available delivery methods to distribute the VoD will rapidly. Finally, the P2P delivery method will be used, only when needed, while at the time that this happens, the number of users already possessing the specific video will be satisfactory with those users, acting as seeders to distribute the VoD to the others.

The aforementioned load balancing algorithm for the load balancing among data centers (DCs) is presented. Algorithm 2 takes as input the table $W[x, y]$ that includes the workload of each VoD on each DC and the number of DCs to be used for the specific content delivery. It updates the table and returns it with the new values to be used for the balanced distribution of the content. The algorithm divides number 1 (the whole percentage) to the number DCs, which will be used, to calculate the portion of delivery requests that each DC should handle. Then, it finds out which column represents the specific VoD, if it already exists, or it assigns a new column for a new VoD. Finally, it selects which DCs will be used, starting with those

Algorithm 1. Delivery Method Selection Algorithm

```

1: procedure SELECTCONTENTDELIVERYMETHODS
2:   neededBW  $\leftarrow$  maximum(currentBW, predictedBW)
3:   epidemicSpread  $\leftarrow$  fits(exponentialFunction)
4:   if epidemicSpread = TRUE then
5:     Use Cloud, all available CDNs and P2P
6:     Reduce the VoD advertisement
7:     Run the load balancing algorithm for DCs
8:   else
9:     switch neededBW do
10:      case neededBW < lowThreshold
11:        Use only Cloud.
12:      case lowThreshold < neededBW <
13:        highThreshold
14:        NoOfDCstoUse  $\leftarrow$ 
15:         $\frac{\text{neededBW} - \text{lowThreshold}}{\text{highThreshold} - \text{lowThreshold}} * \text{AvailDCs}$ 
16:        Use Cloud and NumberOfDCstoUse DCs.
17:        Run the load balancing algorithm for DCs
18:      case neededBW > highThreshold
19:        Use Cloud, all available CDNs and P2P.
20:        Run the load balancing algorithm for DCs

```

Algorithm 2. Load Balancing Algorithm

```

1: procedure LOADBALANCING
2:   portion  $\leftarrow$  1/NoOfDCsToUse
3:   if VoD=new then
4:     Add a column to the  $W[x, y]$  table
5:     column = Column that represents the current VoD
6:     Create a sorted table with the rows,
7:     based on the number of 0 they include
8:      $W[x, \text{column}] = \text{portion}$ ,
9:     where x takes the first NoOfDCsToUse
10:    values of the sorted table

```

that have more zero values in their row, meaning that they do not serve many VoD channels.

IV. PERFORMANCE EVALUATION ANALYSIS AND EXPERIMENTAL RESULTS

This section demonstrates the ability of the epidemic models to describe the spread of content over content delivery systems, and the effectiveness of the RPE to predict future values of network metrics of the utilization of network paths. The effectiveness of the whole proposed system is demonstrated by simulations of the usage scenarios. Finally, this section makes a comparison of the ability of the proposed epidemic model that was specially designed for describing content delivery, to make an early prediction of an upcoming epidemic spread of content compared to other general purpose epidemic models. The monitoring data were collected by a VoD platform [40] and all the videos had a resolution of 480p.

For the evaluation of the forecast algorithms for the short-term prediction, the monitoring data of the bandwidth usage for serving the need of a specific VoD were utilized. The collected

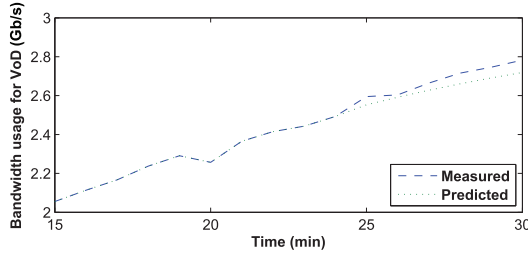


Fig. 7. Measured versus predicted value for the bandwidth usage for VoD.

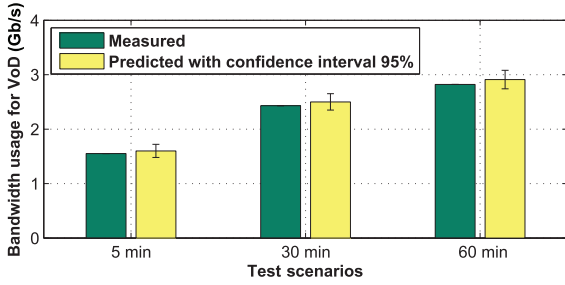


Fig. 8. Test scenarios for bandwidth usage of VoD prediction.

measurements were for a total of 30 min with a period of 5 s, but to avoid periodicity of data, the mean value per minute was used. The 80% (24 min) of the data was exploited to feed the prediction engine. The RPE as described in Section III-D fits the data into the most suitable model and it manages to perform a prediction for the future values. The remaining 20% of the data is then used for the evaluation, through a comparison between the predicted and the actual values as shown in Fig. 7. The important part of the graph is after the first 24 min, where it is clearly depicted that the measured values remain very close to the predicted ones.

The prediction performance of the RPE is presented in Fig. 8. It includes the predicted value showing also the limit of 95% confidence and the corresponding (after the time passes) measured values of bandwidth needs for a VoD channel. The test scenarios presented are for 5, 30, and 60 min prediction. It is clear that the predicted values are near the actual values measured and in all cases the upper and lower limits of the 95% confidence interval include the measured value.

For the evaluation of the long-term prediction models, we collected the number of views per day for two videos provided by a VoD platform [40]. The collected data are for 175 days. The 80% of the data was exploited to feed the prediction engine, while the rest of the data was used for the evaluation through a comparison between the predicted and the actual values as shown in Fig. 9. In the delivery of the first video, as shown in Fig. 9(a), the time-series models are able to predict future demands with enough accuracy. In the second video delivery, as shown in Fig. 9(b), the time-series models are incapable to predict the upcoming enormous increase of the demand since the measured data did not fit well in any time-series model. The prediction based on the time-series models forecasted a steady number of viewers that was proven to be wrong. The prediction based on the epidemic model managed to forecast the upcoming epidemic spread and the predicted values are close to the

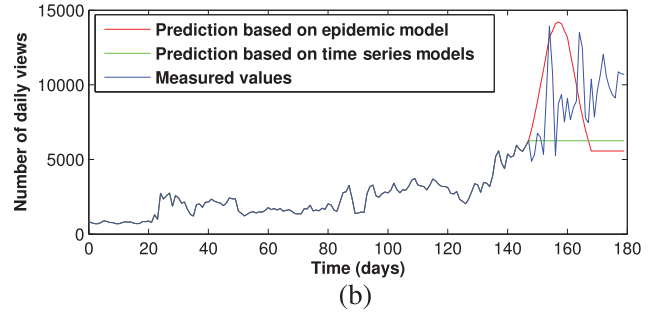
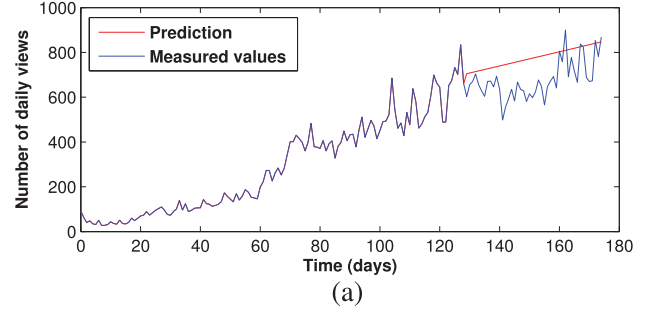


Fig. 9. Measured and predicted number of views per day. (a) Average video. (b) Video that spreads epidemically.

measured values. So, the advantage of using Algorithm 1 is depicted since the algorithm tries to fit the epidemic models and then if there is no epidemic spread it utilizes the time-series models.

For the evaluation of the ability of the proposed epidemic model to make early prediction of an epidemic spread of content through a VoD platform, we compare its performance with that of some general purpose epidemic models. The comparison involves our proposed epidemic model that was specifically designed for describing the spread of content and presented in Section III-C, the basic SIR model as described in detail in Section II and the MSIR model [31]. We collected from a VoD platform history data for 17 videos that we consider as viral because they had a peak demand of more than 10% of the total subscribers as simultaneous viewers. The data about each Video includes measurements about the number of users that were downloading the video each moment, and so we managed to divide them into the basic compartments of the three models. In each model, a number of variables can get values under some constrains. We extended the forecast package of R-system [39] to represent these models and we used a function that automatically assigns values to the variables to fit in the model based on the history data. The amount of data needed in order to deduce that the data can be fitted in the model reveals the time when the actual prediction will be possible in real time. Fig. 10 presents the needed time to predict the epidemic spread, as a percentage of the total lapse time from the release of the video until reaching the maximum value of simultaneous users, for multiple maximum values of simultaneous users based on the seventeen videos of the sample. The graph clearly depicts that all models are able to predict the epidemic spread and in case of a high epidemic spread the prediction is early. In all cases, the proposed model has better results since it makes an earlier prediction. The SIR model does not perform very well

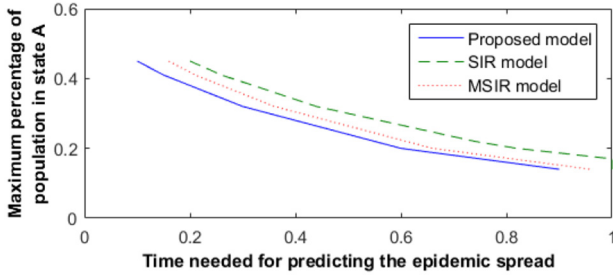


Fig. 10. Time needed for the prediction of the epidemic behavior.

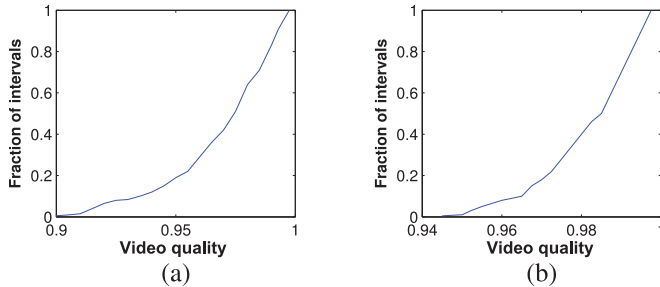


Fig. 11. Fraction of intervals of QoE (a) without and (b) with MDM.

possibly because of the lack of the needed states to represent accurately the content delivery process. MSIR model performs well, possibly because it has the state M (maternally derived immunity) that corresponds to the state T (turn-down) state of the proposed model. It performs worse than the proposed model since it does not distinguish the active users with the infected, nor the recovered to the deleted. The proposed model, provides more variables for customization, and it can be customized to represent the process of content delivery more accurately.

A metric with significant importance in the content delivery process is the QoE for the users. In [3], Yin *et al.* defined quality metrics taking into consideration the time spent in buffering a video compared to the total viewing time to conclude about the QoE. They managed to present results with mean value for the quality metric greater than 0.99 with all measures greater than 0.95. In our approach and since we have a complete network architecture, the MQM component is able to calculate the QoE based on specific dynamical algorithms [41]. The algorithms combine TCP-, buffer-, and media content-related metrics as well as user requirements and expectations to extract a value in the range 0–1 that describes the QoE for each video stream. The closest to 1 implies the highest quality. We collected monitoring results for a specific popular VoD 2 days with normal demand. The first day the MDM was disabled and none of its functionalities were used. As shown in Fig. 11, the fraction confidence interval of QoE for the users when the MDM is not utilized is higher than 0.90 while the mean value is about 0.99. In the second day, when we had the utilization of MDM, the QoE remains over 0.94 for all cases with an average of 0.999.

The rest of the section presents experimental simulation results for evaluation of the performance and the offered reliability in streaming activities, offered by the proposed system.

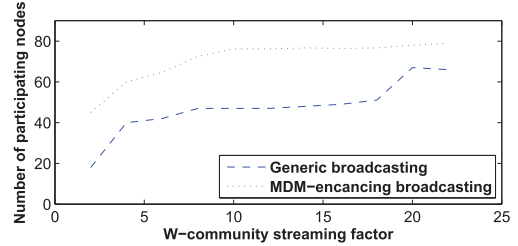


Fig. 12. Number of participating nodes with community streaming factor.

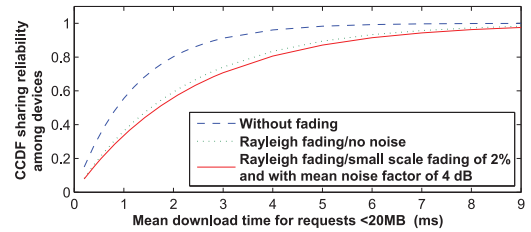


Fig. 13. CCDF sharing reliability among devices with download time (ms) for requests < 20 MB.

Toward implementing such scenario, a common look-up application service for video streaming is set in each node, to enable nodes requesting a stream from a certain user. Fig. 12 shows that the number of the participating nodes is increasing, when MDM-enhancing broadcasting is used, instead of a generic broadcasting. This indicates the enhancement that has been done by the MDM in the broadcasting process, whereas the community streaming factor W , as introduced in [42], indicates the level of robustness in receiving neighboring feedback during the process of streaming. The total delay time with the number of simultaneous transmissions is shown in Fig. 13. The total measured delay is significantly reduced in the presence of MHGC (P2P delivery), whereas the utilization of the existing infrastructure increases the overall delays when multiple transmissions take place.

V. CONCLUSION

This paper presents a novel network architecture, a novel epidemic spread model, and two algorithms for optimal selection of content delivery methods. The proposed epidemic model was analyzed and it was shown through simulation results that it can describe the VoD process when the social interactions among users are high for the specific VoD. Based on the analysis, we were able to perform prediction for the VoD spread based on the model. This paper also presents two algorithms that take advantage of epidemic models prediction and time-series analysis prediction for selecting the optimal delivery method and for load balancing between DCs. The experimental results prove that the prediction engine is accurate and overall the content delivery process gets benefits from the utilization of the model and algorithms. Future directions in our on-going research encompass the further study of the epidemic model for the export of multiple predicted metrics that could be utilized by algorithms for a more accurate forecast of the need or even a better localization of the demand.

REFERENCES

- [1] CNI, CISCO. (2015, Jan. 19). *Consumer Internet Traffic 2012-2018* [Online]. Available: http://www.cisco.com/web/solutions/sp/vni/vni_forecast_highlights
- [2] D. Xu, S. S. Kulkarni, C. Rosenberg, and H.-K. Chai, "Analysis of a CDN-P2P hybrid architecture for cost-effective streaming media distribution," *Multimedia Syst.*, vol. 11, no. 4, pp. 383–399, 2006.
- [3] H. Yin *et al.*, "Design and deployment of a hybrid CDN-P2P system for live video streaming: Experiences with LiveSky," in *Proc. ACM Int. Conf. Multimedia*, 2009, pp. 25–34.
- [4] D. Ciullo, V. Martina, M. Garetto, E. Leonardi, and G. L. Torrisi, "Peer-assisted VoD systems: An efficient modeling framework," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 7, pp. 1852–1863, Jul. 2014.
- [5] U. Mandal, M. F. Habib, S. Zhang, C. Lange, A. Gladisch, and B. Mukherjee, "Adopting hybrid CDN-P2P in IP-over-WDM networks: An energy-efficiency perspective," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 6, no. 3, pp. 303–314, Mar. 2014.
- [6] C. X. Mavromoustakis, A. Bourdena, G. Mastorakis, E. Pallis, and G. Kormentzas, "An energy-aware scheme for efficient spectrum utilization in a 5G mobile cognitive radio network architecture," *Telecommun. Syst.*, vol. 59, no. 1, pp. 63–75, 2015.
- [7] C. X. Mavromoustakis, G. Mastorakis, A. Bourdena, and E. Pallis, "Energy-efficient management using a traffic-oriented routing scheme for cognitive radio networks," *Int. J. Netw. Manage.*, vol. 25, no. 6, pp. 418–434, 2015.
- [8] C. X. Mavromoustakis, G. Kormentzas, G. Mastorakis, A. Bourdena, E. Pallis, and J. Rodrigues, "Context-oriented opportunistic cloud offload processing for energy conservation in wireless devices," in *Proc. Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 24–30.
- [9] C. X. Mavromoustakis, G. Mastorakis, A. Bourdena, and E. Pallis, "Energy-efficient resource sharing using a traffic-oriented routing scheme for cognitive radio networks," *IET Netw.*, vol. 3, no. 1, pp. 54–63, 2014.
- [10] G. Zhang, W. Liu, X. Hei, and W. Cheng, "Unreeling Xunlei Kankan: Understanding hybrid CDN-P2P video-on-demand streaming," *IEEE Trans. Multimedia*, vol. 17, no. 2, pp. 229–242, Feb. 2015.
- [11] Y. Kryftis, C. X. Mavromoustakis, J. M. Batalla, G. Mastorakis, E. Pallis, and G. Skourletopoulos, "Resource usage prediction for optimal and balanced provision of multimedia services," in *Proc. IEEE 19th Int. Workshop Comput. Aided Model. Des. Commun. Links Netw. (CAMAD'14)*, 2014, pp. 255–259.
- [12] Y. Kryftis, G. Mastorakis, C. X. Mavromoustakis, J. M. Batalla, A. Bourdena, and E. Pallis, "A Resource Prediction Engine for Efficient Multimedia Services Provision," in *Resource Management of Mobile Cloud Computing Networks and Environments*. Hershey, PA, USA: IGI Global, pp. 361–380, 2015.
- [13] Y. Kryftis *et al.*, "Resource usage prediction algorithms for optimal selection of multimedia content delivery methods," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2015, pp. 5903–5909.
- [14] Y. Kryftis, C. X. Mavromoustakis, G. Mastorakis, J. M. Batalla, and P. Chatzimisios, "Epidemic models using resource prediction mechanism for optimal provision of multimedia services," in *Proc. IEEE 20th Int. Workshop Comput. Aided Modell. Des. Commun. Links Netw. (CAMAD'15)*, 2015, pp. 91–96.
- [15] Y. Kryftis, G. Mastorakis, C. X. Mavromoustakis, J. M. Batalla, E. Pallis, and G. Kormentzas, "Efficient entertainment services provision over a novel network architecture," *IEEE Wireless Commun. Mag.*, vol. 23, no. 1, pp. 14–21, Feb. 2016.
- [16] D. Niu, Z. Liu, B. Li, and S. Zhao, "Demand forecast and performance prediction in peer-assisted on-demand streaming systems," in *Proc. IEEE INFOCOM*, 2011, pp. 421–425.
- [17] G. E. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*. Hoboken, NJ, USA: Wiley, 2013.
- [18] D. Niu, H. Xu, B. Li, and S. Zhao, "Quality-assured cloud bandwidth auto-scaling for video-on-demand applications," in *Proc. IEEE INFOCOM*, 2012, pp. 460–468.
- [19] V. Assimakopoulos and K. Nikolopoulos, "The theta model: A decomposition approach to forecasting," *Int. J. Forecast.*, vol. 16, no. 4, pp. 521–530, 2000.
- [20] R. J. Hyndman, M. L. King, I. Pitrun, and B. Billah, "Local linear forecasts using cubic smoothing splines," *Aust. N. Z. J. Stat.*, vol. 47, no. 1, pp. 87–99, 2005.
- [21] B. Klaseen, "Efficient content distribution in social-aware hybrid networks," *J. Comput. Sci.*, vol. 4, no. 4, pp. 209–218, 2013.
- [22] G. Skourletopoulos, C. X. Mavromoustakis, G. Mastorakis, J. J. P. C. Rodrigues, P. Chatzimisios, and J. M. Batalla, "A fluctuation-based modelling approach to quantification of the technical debt on mobile cloud-based service level," in *Proc. IEEE GLOBECOM, Int. 12 Workshop Cloud Comput. Syst. Netw. Appl. (CCSNA)*, 2015, pp. 1–6.
- [23] A. Bourdena, C. Mavromoustakis, G. Mastorakis, J. Rodrigues, and C. Dobre, "Using socio-spatial context in mobile cloud offload process for energy conservation in wireless devices," *Trans. Cloud Comput.*, 2015.
- [24] P. Gonçalves, R. Shubhabrata, T. Begin, and P. Loiseau, "Dynamic resource management in clouds: A probabilistic approach," *IEICE Trans. Commun.*, vol. 95, no. 8, pp. 2522–2529, 2012.
- [25] M. J. Keeling and K. T. Eames, "Networks and epidemic models," *J. R. Soc. Interface*, vol. 2, no. 4, pp. 295–307, 2005.
- [26] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Phys. Rev. Lett.*, vol. 86, no. 14, p. 3200, 2001.
- [27] W. O. Kermack and A. G. McKendrick, "Contributions to the mathematical theory of epidemics. II. The problem of endemicity," *Proc. R. Soc. London A*, vol. 138, no. 834, pp. 55–83, 1932.
- [28] N. T. J. Bailey, *The Mathematical Theory of Epidemics*. London: Charles Griffin & Company Limited, 1957.
- [29] R. M. Anderson and R. M. May, *Infectious Diseases of Humans*. London, U.K.: Oxford Univ. Press, 1991.
- [30] P. Mafuta, J. Mushanyu, and G. Nhawu, "Extension of the analysis of the MISR model by applying the Laplace-adomian decomposition method," *World J. Modell. Simul.*, vol. 10, no. 3, pp. 185–192, 2014.
- [31] D. Bichara, A. Igdir, and G. Sallet, "Global analysis of multi-strains SIS, SIR and MSIR epidemic models," *J. Appl. Math. Comput.*, vol. 44, nos. 1–2, pp. 273–292, 2014.
- [32] S. Meloni, A. Arenas, and Y. Moreno, "Traffic-driven epidemic spreading in finite-size scale-free networks," *Proc. Nat. Acad. Sci.*, vol. 106, no. 40, pp. 16 897–16 902, 2009.
- [33] C. X. Mavromoustakis and H. D. Karatza, "A gossip-based optimistic replication for efficient delay-sensitive streaming using an interactive middleware support system," *IEEE Syst. J.*, vol. 4, no. 2, pp. 253–261, Jun. 2010.
- [34] B. K. Mishra and N. Jha, "Fixed period of temporary immunity after run of anti-malicious software on computer nodes," *Appl. Math. Comput.*, vol. 190, no. 2, pp. 1207–1212, 2007.
- [35] C. X. Mavromoustakis and H. D. Karatza, "Segmented file sharing with recursive epidemic placement policy for reliability in mobile peer-to-peer devices," in *Proc. IEEE 13th Int. Symp. Model. Anal. Simul. Comput. Telecommun. Syst.*, 2005, pp. 371–378.
- [36] JRI. (2015, Jan. 19). rJava - Low-level R to Java interface [Online]. Available: <http://rforge.net/rJava/index.html>
- [37] R. C. Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2014 [Online]. Available: <http://www.R-project.org/>.
- [38] Forecast Package. (2015, Jan. 19). *R-Project* [Online]. Available: <http://CRAN.R-project.org/package=forecasting>
- [39] R. J. Hyndman and Y. Khandakar, "Automatic time series for forecasting: The forecast package for R," Dept. Econ. Bus. Stat., Monash Univ., Melbourne, Vic., Australia, Tech. Rep. 06/07, 2007.
- [40] Cablenet. (2015). Video on demand [Online]. Available: <http://cablenet.com.cy/vod/>
- [41] C. Alberti *et al.*, "Automated QoE evaluation of dynamic adaptive streaming over HTTP," in *Proc. 5th Int. Workshop Qual. Multimedia Experience*, 2013, pp. 58–63.
- [42] C. X. Mavromoustakis and H. D. Karatza, "Performance evaluation of opportunistic resource-sharing scheme using socially oriented outsourcing in wireless devices," *Comput. J.*, vol. 56, pp. 184–197, 2012.



Yiannos Kryftis (M'16) received the B.Sc. degree in electrical and computer engineering from the National Technical University of Athens, Athens, Greece, in 2008 and the M.Sc. degree in advanced computing from Imperial College, London, U.K., in 2009. He is currently pursuing the Ph.D. degree in management architectures and systems for future internet networks at the National Technical University of Athens.

Currently, he is a Researcher with the Department of Computer Science, University of Nicosia, Nicosia, Cyprus. His research interests include policy-based network management, network architectures for the future Internet, virtualization information models and algorithms in federated environments, and multimedia services provisioning.



George Mastorakis (M'12) received the B.Eng. (Hons.) degree in electronic engineering from the University of Manchester Institute of Science and Technology (UMIST), Manchester, U.K., in 2000; the M.Sc. degree in telecommunications from the University College London (UCL), London, U.K., in 2001; and the Ph.D. degree in telecommunications from the University of the Aegean, Mytilene, Greece, in 2008.

He is serving as an Assistant Professor with the Technological Educational Institute of Crete, Heraklion, Greece, and a Research Associate with the Research and Development of Telecommunications Systems Laboratory, Centre for Technological Research of Crete, Heraklion, Greece. He has a more than 100 publications at various international conferences proceedings, workshops, scientific journals, and book chapters. His research interests include cognitive radio networks, networking traffic analysis, radio resource management, and energy-efficient networks.



Constandinos X. Mavromoustakis (M'04) received a five-year Dipl.Eng. degree (B.Sc., B.Eng., M.Eng.) in electronic and computer engineering from the Technical University of Crete, Chania, Greece, in 2000, the M.Sc. degree in telecommunications from the University College of London, London, U.K., in 2001, and the Ph.D. degree in informatics from Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2006.

He is currently an Associate Professor with the Department of Computer Science, University of Nicosia, Nicosia, Cyprus. He is leading the Mobile Systems Laboratory (MOSys Lab., <http://www.mosys.unic.ac.cy/>), Department of Computer Science, University of Nicosia, dealing with design and implementation of hybrid wireless testbed environments and MP2P systems, IoT configurations and smart applications, as well as high-performance cloud and mobile cloud computing (MCC) systems, modeling and simulation of mobile computing environments and protocol development and deployment for large-scale heterogeneous networks and new "green" mobility-based protocols. He has a dense research work outcome in distributed mobile systems and spatio-temporal scheduling, consisting of numerous refereed publications including several Books (IDEA/IGI, Springer and Elsevier).

Dr. Mavromoustakis has been an active member (vice-chair) of the IEEE/R8 regional Cyprus section since January 2016, and since May 2009, he has been as the Chair of C16 Computer Society Chapter of the Cyprus IEEE section. He has served as a Consultant to many industrial bodies [i.e., member of the Technical Experts for Internet of Things-IoT competition at Intel Corporation LLC (www.intel.com) for the ChallengeMe etc.]. He is a management member of the IEEE Communications Society (ComSoc) Radio Communications Committee (RCC) and a board member the IEEE-SA Standards IEEE SCC42 WG2040, whereas he has served as track Chair and Co-Chair of various IEEE International Conferences (including AINA, IWCMC, ICC, GlobeCom, and IEEE Internet of Things).



Jordi Mongay Batalla (M'15) received the M.Sc. degree in from the Universitat Politècnica de València, Valencia, Spain, in 2000, and the Ph.D. degree from Warsaw University of Technology, Warsaw, Poland, in 2009, both in telecommunications.

He is an Assistant Professor with Warsaw University of Technology. He worked with Telcordia Poland (Ericsson R&D), Poznan, Poland, where he has been with the National Institute of Telecommunications since 2010, and the Head of the Internet Architectures and Applications Department. He has authored or coauthored more than 100 papers published in books, and international and national journals and conference proceedings and an Editor member of several journals and conferences. His research interests include quality of service/quality of experience in IPv4/v6 and Software Defined Networks (SDN) infrastructures, especially multimedia adaptive streaming, future Internet architectures [content-aware networks, information-centric networks, architectures for Network Function Virtualization (NFV)] as well as applications for future Internet [Internet of Things, smart cities, Internet Protocol Television (IPTV)].



Joel J. P. C. Rodrigues (SM'06) received the B.Sc. degree (licentiate) in informatics engineering from the University of Coimbra, Coimbra, Portugal, in 1995; the M.Sc. degree in information and knowledge management and Ph.D. degree in informatics engineering from the University of Beira Interior, Covilhã, Portugal, in 2002 and 2006, respectively; and the Habilitation degree in computer science and engineering from the University of Haute Alsace, Mulhouse, France, in 2014.

He is a Professor with the Department of Informatics, University of Beira Interior, Covilhã, Portugal, and a Researcher with the Instituto de Telecomunicações, Lisboa, Portugal. He has authored or coauthored more than 400 papers in refereed international journals and conferences, and a book, and holds three patents. His research interests include sensor networks, e-health, e-learning, vehicular delay-tolerant networks, and mobile and ubiquitous computing.

Dr. Rodrigues is a Licensed Professional Engineer (as a senior member), a member of the Internet Society, an International Academy, Research, and Industry Association (IARIA) Fellow, and a senior member of Instituto de Telecomunicações (IT) and Association for Computing Machinery (ACM). He is the Leader of Next Generation Networks and Applications Group (NetGNA) Research Group (<http://netgna.it.ubi.pt>), the Chair of the IEEE ComSoc Technical Committee on eHealth, the Steering Committee Member of the IEEE Life Sciences Technical Community, a Member Representative of the IEEE Communications Society on the IEEE Biometrics Council, and the Officer of the IEEE 1907.1 standard. He is the Editor-in-Chief of the *International Journal on E-Health and Medical Communications*, the *Recent Advances on Communications and Networking Technology*, and the *Journal of Multimedia Information Systems*, and an Editorial Board Member of several journals. He has been a General chair and a TPC Chair of many international conferences, including the IEEE International Conference on Communications (ICC) and IEEE Global Communications Conference (GLOBECOM). He is a member of many international TPCs and participated in several international conferences organization. He had been awarded several Outstanding Leadership and Outstanding Service Awards by the IEEE Communications Society and several best papers awards.



Ciprian Dobre (M'06) received the Ph.D. degree in computer science from the University Politehnica of Bucharest, Bucharest, Romania, in 2008.

He has scientific and scholarly contributions in the field of large-scale distributed systems concerning mobile applications and smart technologies to reduce urban congestion and air pollution Models and Techniques for Optimising Traffic in Urban Environments (TRANSYS), context-aware applications Context-Aware Platform using Integrated Mobile Services (CAPIM), opportunistic networks and mobile data offloading [Social Prediction-Based Opportunistic Routing (SPRINT), Collaborative Selfish Node Detection and Incentive Mechanism for Opportunistic Networks (SENSE)], MONitoring Agents using a Large Integrated Services Architecture (MonALISA), high-speed networking [Virtual Intelligent Networks for Computing Infrastructures (VINCI), Fast Data Transfer (FDT)], grid application development [Enabling Grids for E-science (EGEE), South Eastern European Grid (SEE-GRID)], and evaluation using modeling and simulation [Models of Networked Analysis at Regional Centers (MONARC 2), Vehicular Networks Simulator (VNSim)]. These contributions led to important results, demonstrating his qualifications and potential to go significantly beyond the state of the art.

Dr. Dobre was awarded a Ph.D. scholarship from California Institute of Technology (Caltech, CA, USA), and another one from Oracle. His results received one IBM Faculty Award, two Corporation for Education Network Initiatives in California (CENIC) Awards, and three Best Paper Awards (in 2013, 2012, and 2010). The results were published in 6 books, 10 chapters in edited books, 34 papers in major international peer-reviewed journal (19 as main author, cumulated impact factor of 13.24), over 100 papers in well-established international conferences and workshops (with over 200 citations).