Yang, P, Fan, Q, Yin, H, Min, G, Luo, Y, Lyu, Y, Huang, H and Jiao, L

Video Delivery Networks: Challenges, Solutions and Future Directions

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Yang, P, Fan, Q, Yin, H, Min, G, Luo, Y, Lyu, Y, Huang, H and Jiao, L (2017) Video Delivery Networks: Challenges, Solutions and Future Directions. Computers and Electrical Engineering, 66. pp. 332-341. ISSN 0045-7906

Video Delivery Networks: Challenges, Solutions and Future Directions

Qilin Fan[1], Hao Yin[1], Geyong Min[2], Po Yang[3], Yan Luo[4], Yongqiang Lyu[1], Haojun Huang[5] and Libo Jiao[1]

[1]Research Institute of Information Technology (RIIT), Tsinghua University

[2]Department of Mathematics and Computer Science, University of Exeter

[3]Department of Computer Science, Liverpool John Moores University

[4]Department of Electrical and Computer Engineering, University of Massachusetts Lowell

[5]Department of Communication Engineering, Wuhan University

ABSTRACT

Internet video ecosystems are faced with the increasing requirements in versatile applications, ubiquitous consumption and freedom of creation and sharing, in which the user experience for high-quality services has become more and more important. Internet is also under tremendous pressure due to the exponential growth in video consumption. Video providers have been using content delivery networks (CDNs) to deliver high-quality video services. However, the new features in video generation and consumption require CDN to address the scalability, quality of service and flexibility challenges. As a result, we need to rethink future CDN for sustainable video delivery. To this end, we give an overview for the Internet video ecosystem evolution. We survey the existing video delivery solutions from the perspective of economic relationships, algorithms, mechanisms and architectures. At the end of the article, we propose a data-driven information plane for video delivery network as the future direction and discuss two case studies to demonstrate its necessity.

Keywords

Video Delivery Network, Scalability, Quality of Service, Cloud computing, Data-Driven, Information Plane

1. INTRODUCTION

Video is a kind of important media for communication and entertainment. In the past decade, the consumption of video contents changed from offline viewing to online streaming, which imposed stringent requirements on the network latency and throughput of video delivery network. Internet video ecosystem has experienced new characteristics accordingly: (1) the diversity of video applications, such as online streaming (e.g., IPTV and Netflix), and real-time video conferencing (e.g., Skype), (2) the ubiquitous video consumption by users who watch videos wherever they are and whenever they can; and (3) the freedom in the creating, distributing and sharing of user generated video contents through online video services such as YouTube, Vimeo or Hulu. The situation is further deteriorated by the emerging trends of adopting higher definition video streams, requesting more and more bandwidth. The Cisco predicts that video consumption will amount to 82% of the global consumer traffic in 2020 [1]. In conjunction with these growing traffic volumes, users' expectations of high quality of experience (e.g., high-definition video, low re-buffering, low startup delays) are continuously increasing [2].

Content Delivery Networks (CDNs) have been playing a critical role in content hosting, orchestration, mashup, transport, and edge access. Today, video providers rely on CDNs, such as Akamai and Limelight to serve video content across different geographical locations. However, with the new characteristics of Internet video generation and consumption, the architecture of today's HTTP-based CDN is being stressed [3].

As a result, there is a strong call for rethinking future evolution for sustainable video delivery. The state of the art in video delivery network has experienced a major boost in recent years, measured in the number of publications and industrial groups focused on the topic. Diverse interests have led to heterogeneous, fragmented landscapes. In this article, we give a comprehensive review and analysis on video delivery network, focusing on the major obstacles in the way of further progress and suggesting possible future direction.

The rest of the article is organized as follows. We firstly give an overview of both the evolution of Internet video ecosystem and the key challenges to video delivery network in new personalization phase. After a brief summary on the surveys in this field, we illustrate how existing approaches work to address parts of challenges from the perspective of economic relationships, algorithms, mechanisms and architectures. Finally, we suggest constructing data-driven multidimensional information plane for video delivery network as future research trend and conclude the article.

2. VIDEO DELIVERY NETWORK: EVOLUTION AND CHALLENGES

Figure 1 shows the evolution of video ecosystem experienced several important phases in the last two decades.

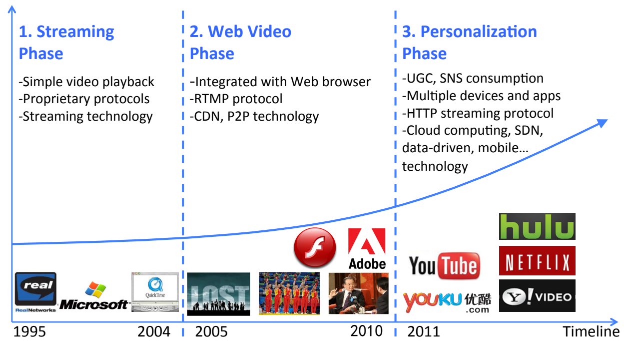

Figure 1: Evolution of Internet video ecosystem.

In the mid-1990s, the Internet boosted from a text-based system to a multimedia-contained system. The representative streaming technique changed the way videos that are transmitted over the Internet, from the time-consuming downloads to real-time viewing. In 2005, Adobe system acquired the flash player which was originally developed by Macromedia, and flash soon became the de-facto standard for web-based streaming video (over RTMP) in this wave when videos explosively transmitted across the Internet along with widespread broadband access. Meanwhile, CDN and Peer-to-Peer (P2P) technologies emerged with the tide of the times.

Since 2011, the portable devices with cameras have become more and more ubiquitous, which enable users conveniently create and share videos. Users are increasingly demanding content on their terms - any content, any time, any device, any place. In addition, another key driver to the rapid explosion of streaming video was the technical shift from specialized streaming protocols such as RTMP to HTTP chunk-based streaming protocols. The use of commodity service greatly decreases the barrier of access entry by allowing video providers to utilize the existing HTTP-based CDN infrastructures in delivering videos to wide audiences. New applications, devices and protocols not only promote the Internet video ecosystem in the personalization phase, but also bring new challenges to the video delivery network. The key challenges can be summarized as follows.

1) **High scalability**: We are witnessing a proliferation of videos over the Internet. For example, Netflix, a web service that streams premium video contents, accounted for 37 percent of peak download Internet traffic in North America in 2015. As of July 2015, more than 400 hours of video were uploaded to YouTube every minute, up from 300 hours per minute in November 2014. According to Cisco report [1], IP video traffic will be 82% of all consumer Internet traffic by 2020, up from 70% in 2015. Today, boosted by the strong growth of newly uploaded videos and demand for various bitrate, storage and network bandwidth become serious bottlenecks for large scale video delivery.

2) **High Quality of Service (QoS)**: Video contents are now being increasingly consumed on big screens. Users expect much higher quality of video service than ever. The existing study [2] showed that 1% increase in buffering ratio reduces user engagement by 3 minutes in 2010 but 6 minutes in 2012. Therefore, video delivery service providers must keep improving server-side and network-level performance for better QoS.

3) **High flexibility**: Video consumption has shifted dramatically from a relatively predictive and controlled model to a highly dynamic mode, such as the large traffic spikes during online live and on-demand events. Streaming large live videos and popular on-demand titles can cause significant network congestion leading to poor subscriber experience. Traditional CDNs, however, are statically deployed, and their centralized redirection mechanisms are not good for dynamic adjustment, which results in performance penalties. On the other hand, traditional CDNs were horizontal - designed to deliver all forms of content by the same mechanism across all user categories. The flexible service depending on different business models require new CDNs that should be vertically customized according to different content and user requirements.

3. EXISTING SOLUTIONS

To address the aforementioned challenges, a range of approaches were proposed in recent years. We survey the existing methods and classify them into four categories according to the challenge type the approach focused on and the perspective from which the approach solved (shown in table 1).

3.1 Cooperation between stakeholders

The stakeholders in video delivery network ecosystem can be grouped into the client, CDN, Internet Service Provider (ISP), and Content Provider (CP). CP has a central database of content and pays CDN for content delivery service. The CDN is an overlay network that offers additional value to the basic Internet content delivery. It pays for transit service provided by ISP. Client pays ISP for network access and pays CP for watching videos as well. Traditional CDNs do not always have enough service points in locations that can provide good performance to satisfy the growing demands. To address such a scalability issue, the boundaries among different stakeholders in video delivery networks are blurring for cooperation. An overview of the involvement of content delivery stakeholders is shown in figure 2.

Driven by economic policies, several new hybrid architectures were proposed:

1) CPs deploy application-specific CDNs with direct peering with or inside ISPs, e.g., Netflix's platform - Open Connect for video stream delivery and Google Global Cache, primarily for YouTube. It allows ISPs to cache the content locally, and thus reduces the strain of CP's traffic over the network.

2) Commercial CDNs and clients operate hybrid delivery systems where clients download contents from servers as well as from other clients. CDNs provide excellent quality when workloads are within the provisioning limits; otherwise, P2P technologies can be used to solve the scalability issue by utilizing the resources of the

Table 1: **The categories of the existing solutions.**

What does it focus on?	From which perspective?	How does it solve?
Scalability	Economic relationship adjustment	Cooperation between stakeholders (Sec. 3.1)
QoS	Algorithm optimization	Bitrate and CDN/server adaption (Sec. 3.2)
Flexibility	Infrastructure as a service	Cloud computing utilization (Sec. 3.3)
Scalability	Architecture redesign	Information-centric network (Sec. 3.4)

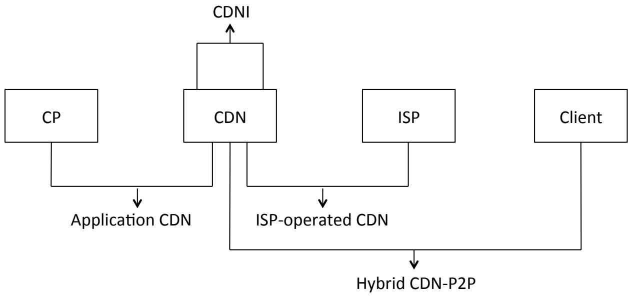

Figure 2: Involvement of content delivery stakeholders.

participating peers. This hybrid CDN-P2P architecture has been widely used in live streaming and social video propagation[4].
3) ISPs, on one hand, are vulnerable to the traffic dynamics caused by time-varying shifts in content demand. On the other hand, they wish to take advantage of their regional presence and proximity to end users to enter the content delivery market [5]. This leads many ISPs to deploy proprietary CDNs within their networks, which allows them to obtain extra income from CP by offering their CDN services to them as a complement to pure-play CDNs.
4) A further-reaching solution was also discussed when Cisco began to explore the feasibility of CDN federations since 2011. CDNs in federation are supposed to be interconnected via open interfaces and shared by autonomous members so they can act as a multi-footprint logical CDN. Internet Engineering Task Force (IETF) working group was also established to define the standards of CDN interconnection (CDNI) [6].

The alliances between stakeholders show a natural evolution of video delivery network today, which was driven by the requirements for storage and bandwidth issues. However, such alliances require meticulous mechanism design and management to guarantee the fairness among the stakeholders.

3.2 Bitrate and CDN/Sever Adaption

The research on the impact of service quality [2] showed that users are quite sensitive to buffering and high startup latency and they prefer to higher bitrate videos. It also identifies that many of the user experience issues today are caused by sub-optimal client-side control logic, and spatial and temporal diversities of performance over different CDNs and networks.

The design space of adaptive algorithms for video delivery quality in recent work can be decomposed into three dimensions [3]:
1) *What parameters can we control?* Bitrate or CDN/server;
2) *Who decides the values for tuned parameters?* Client-side, server-driven or control plane;
3) *When can we choose the parameters?* Startup or midstream.

Table 2 illustrates several representative work in this field by combining different options from these three dimensions.

Clients are very sensitive to local network dynamics. At the same time, they cannot respond efficiently to the network variations due to their distributed nature. Therefore, most bitrate adaptive algorithms are driven by client-side [7][8][9][10]. Nevertheless, the advantage of server-side bitrate adaptive algorithms lies in not requiring clients' cooperation and solving the problem of unfairness and instability [11][12]. There are also some other studies that uncover proprietary CDN/server selection algorithms [13][14]. A client in Netflix adapts to bandwidth variations by adjusting the bitrates and continues to use the same CDN server only when the current CDN server is unable to serve the lowest bitrate, it switches to a different CDN server. Inspired by industrial CDN multihoming applications, Liu *et al.* [15] optimized the CDN selection to make a client actively use additional servers when the service of a single CDN server is insufficient. Study [3][16] even designed a video control plane that can use a global view of CDN and network performance to dynamically assign clients suitable CDNs and bitrates.

The motivation of bitrate adaptive algorithms or CDN/server selection is to deliver as high quality video as possible even in dynamic environment. However, the algorithms use measurement-driven performance feedback without knowledge of bottleneck from the network level. Therefore, there appears to be some more space for performance improvement with network-level information from ISP given.

3.3 Cloud Computing Utilization

To meet volatile user demands for videos, cloud computing emerges as a new paradigm to reshape the pattern of video distribution over the Internet, in which resources (such as CPU, bandwidth, and storage) can be rented on demand from cloud data centers. In cloud computing environment, hardware and software resources are abstracted using virtualization technology for providing service to consumers.

Industrial practitioners have made some efforts to migrate VoD applications to the cloud. Netflix, a major

Table 2: Representative work on adaptive algorithms for video delivery quality.

What parameter?	Who choose?	When to choose?	Related work
Bitrate	Client	Midstream	[7][8][9][10]
Bitrate	Server	Midstream	[11][12]
CDN/Server	Control plane	Startup	[13]
CDN/Server	Control plane + Client	Midstream	[15]
Bitrate CDN/Server	Client Control plane	Midstream Midstream	[14]
Bitrate CDN/Server	Control plane Control plane	Midstream Midstream	[3][16]

VoD provider in North America, has moved its streaming servers, encoding software, search engines, and huge data stores to Amazon Web Services (AWS). Netflix's successful practice verifies that cloud computing service could provide elastic resources even for large-scale video service.

There have also been many studies on utilizing cloud computing resources in academic circles:
1) With intensive and dynamic bandwidth/storage demands in real time, one fundamental question is: *how to quantify dynamic user demands in service delivery, or more precisely, how should a video service provider learn the dynamic demands from users?* Wu *et al.* [17] introduced a Jackson queueing network based model to characterize the viewing behaviors of users inside each channel in the VoD system and studied the upload capacity provided by cloud resources in C/S and P2P mode.
2) Video delivery service is a network-bound service with stringent bandwidth requirements. Therefore, bandwidth reservation becomes a value-added feature offered by cloud services to attract customers with bandwidth-intensive applications. It can be translated into the question: *how should a video service provider dynamically book bandwidth resources from multiple locations and direct user requests to guarantee user's experience?* He *et al.* [18] applied the Nash bargaining solution to solve the joint optimization problem that takes both the operational cost and user experience into account and derived the optimal bandwidth provisioning strategy.
3) In addition to the bandwidth requirement, the storage demand is also significant, especially for UGC or social video services which easily ignite viral propagation effects. Federation of geo-distributed cloud services provides a more economic motivation: "Infinite" on demand resources. Its realize presents challenge on: *how to efficiently store and migrate video contents across different cloud sites?* Wu *et al.* [19] designed an online algorithm that can utilize cloud storage resources to accommodate dynamic demand on the fly and further achieve the optimality via a $\Delta(t)$-step look ahead mechanism.

Moving the whole video delivery to the cloud can be a real challenge as cloud will not replace all traditional hosting or on-premise deployments but complement them. As a result, there will always be situations where security requirements, flexibility, performance or control will preclude the cloud.

3.4 Information-Centric Networking

Since the current growth rate for the bandwidth demand of video applications is difficult to sustain for the network operators, new architecture - information-centric networking (ICN) has been proposed. It allows the network to cache data by attaching storage to routers, eliminating complex optimizations required by CDNs, so that content can be distributed in a cost-efficient and secure manner. This is rooted by the observation that Internet traffic consumption is largely content driven.

Recent studies extended ICN to support video streaming. Video streaming over ICN has faced two challenges with respect to performance. First, Adaptive Bitrate Streaming (ABS) is incompatible with ICN's fundamental principles. For instance, an ABS connection does not support swarming which is natural in ICN. The situation becomes more complicated when multiple users requesting the same content but with different access bandwidth. Yu *et al.* [20] proposed a DASH-aware scheduling algorithm that the network (by way of logically centralized controller) computes the available link resources and schedules the chunk dissemination to edge caches ahead of the user's requests by utilizing the DASH manifest. Second, video traffic volume inflates caching resource requirements. Therefore, caching redundancy in in-network level affects more severely real-time video transfers. Sun *et al.* [21] conducted large-scale trace-driven evaluation and analysis of ICN-on-video and observed that the benefits of ICN for video workloads with caches smaller than 100GB are marginal and discussed that ubiquitous caching is not fundamentally necessary. Therefore, it suggests caching only in specific locations along the content delivery path or using cooperative caching to reduce redundancy.

Although the content-driven principle in ICN is best suit for bandwidth-intensive video applications, from the ISP's point of view, ICN clean slate approaches require intrusive modifications to the communication stack and have yet to demonstrate traffic engineering capabilities. Moreover, addressing, naming and routing issues are not yet solved to scale and need more research.

4. DATA-DRIVEN INFORMATION PLANE

Based on the review of existing approaches, we have a further inspection on the root causes of the challenges. Current layered architecture isolates information ex-

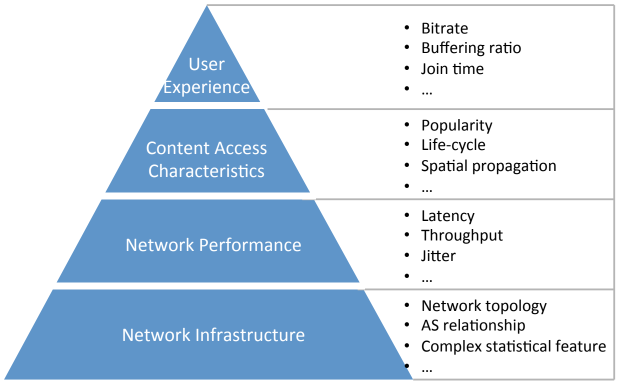

Figure 3: Information plane for video delivery network.

changes between video applications and networks. For video applications, the network states are invisible and network resources are uncontrollable. This also happens as network is unaware of application content and user profile. Therefore, we would like to introduce a data-driven information plane for intermediate layer - video delivery network as a possible future direction. The data-driven information plane consists of four layers from bottom up in figure 3. They can be divided into two parts. Network infrastructure and network performance is related to network providers. Network infrastructure layer contains network structural knowledge such as topology, AS relationships and complex statistical features. Network performance layer reflects network quality of service such as latency, throughput and jitter, which is built upon network infrastructure. Content access characteristics and user experience is related to content providers. Content access characteristics like popularity, life-cycle and spatial propagation shows specific user behaviors in different content providers. User experience layer reflects quality of service provided by content providers such as bitrate, buffering ratio and join time. Based on this information plane, video delivery network can be significantly improved via sophisticated processing such as control, fault diagnosis, scheduling, and decision making. We will take network performance and content access characteristics below as two case studies to demonstrate how information plane plays an important role.

4.1 Network Performance

An up-to-date annotated map of the network performance is significant for network management, task scheduling, and resource optimization in video delivery network. Progress has been made in data models and tools for at-scale network measurements using minimum deployment cost. A network coordinate system [22] models the Internet as a geometric space and characterizes the position of any node in the Internet by a coordinate in this space. IPlane [23] utilizes the Internets routing topology to build an atlas of the Internet, and estimates the properties of the paths between arbitrary end hosts on the Internet based on the path composition technique.

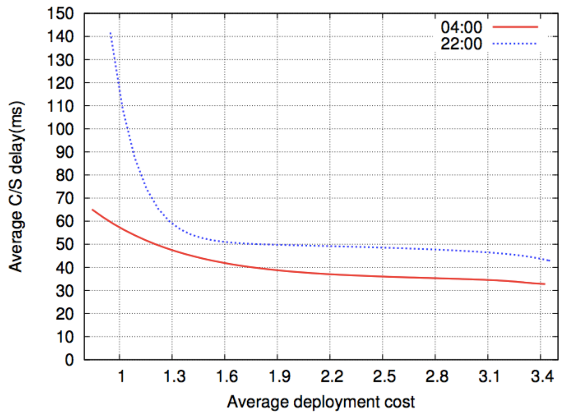

Figure 4: Correlation between average delay and average deployment cost at different time.

We describe our servers placement framework - Netclust [24] as example. Traditional servers placement strategy is a classical facility location problem with the aim of choosing M replicas among N potential sites. Known as a NP-hard problem, it does not scale well with the number of clients and the number of server candidates with the rapid growth of cloud computing platforms. It also fails in discovering the inherent rule and potential trends of the service performance as the number of servers to be deployed changes, and thus cannot give insights on the choice of an ideal number of servers. Given large-scale network performance in information plane, Netclust could leverage a clustering algorithm to determine the correlation between the average service delay and the unit deployment cost for busy period and idle period. Figure 4 shows such relationship. Firstly, we observe that the service delay decreases as deployment cost increases, following the similar negative exponential distribution. Secondly, the delay-cost curve of “busy network” is converging faster than “idle network”. Moreover, with the same deployment cost, “idle network” has the better performance than “busy network”. Operators can easily make the most suitable decision based on the relationship, analyzed above, between the number of servers, deployment cost and service performance.

4.2 Content Access Characteristics

With regard to replication and caching mechanisms in video delivery networks, it is natural to answer which videos should be replicated, how many replicas will be, and where they will reside. Take video on demand for user generated content as an example, the approaches can be divided into two classes: (1) Proactive approach: traditional replica placement, together with request routing is formulated as an optimization problem given global prior information such as contents popularity and network congestion. However, its high computing overhead and lack of quick response to access interest changes limit its practical application. (2) Reactive approach: the contents for caching are driven by users’ requests. However, its video popularity blindness characteristic

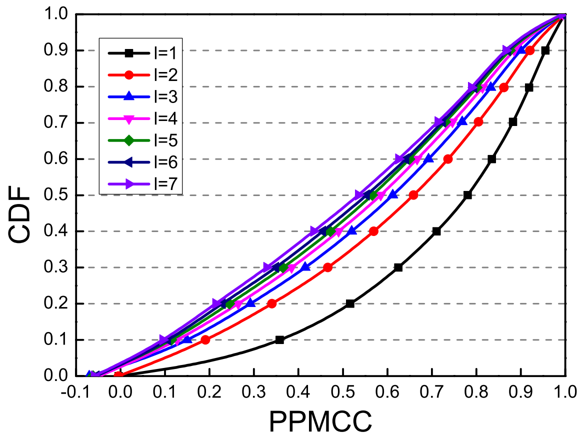

Figure 5: Cumulative distribution of videos' PPMCC for different temporal lag.

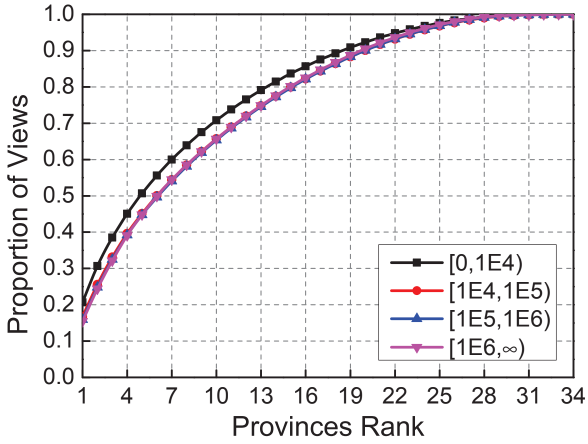

Figure 6: Geographic cumulative distribution of views for videos with different popularity.

leads to cache pollution, where the cache is wasted by low demand videos, resulting in low hit rate and high bandwidth cost. We have worked on understanding the properties of videos and how the users access behaviours can be modelled and predicted in our recent research. We plot cumulative distribution function of pearson product-moment correlation coefficient (PPMCC) for different temporal lags scaling from 1 to 7 for sampled videos in figure 5. We observed that the median PPMCC for videos of temporal lag 1 is 0.78, which shows strong correlation from temporal perspective. For temporal lags scaling from 2 to 7, the median PPMCCs are still larger than 0.5 but decrease slowly as the temporal lag grows. It suggests that videos popularity is predictable using recent access times and recent 4 days values are enough for predicting. We also examined the geographic distribution of views for videos in province granularity. The interval in each colored line denotes video access times. Point (x,y) in the figure 6 means that the x top locations for videos account for y% of the total number of views in each category. It illustrates that 30% provinces could cover 70% of video views for all categories. This geographic locality feature illustrates that content replicating in "appropriate small-scale locations can achieve most of the traffic. Research [4] also extract the content access characteristics of social videos from correlation analysis at information plane to guide propagation-aware replication and improve video service quality.

4.3 Complementary

There are also several works in the recent literature dealing with network infrastructure aware or user experience aware video delivery network. Studies [25][26] developed algorithms to heuristically infer AS relationships based on AS path information available in public BGP data. Accurate and complete AS relationships could provide a scientific basis for decision-making of cost-effective cache deployment for CDN [27]. Balachandran *et al.* [28] collected statistics directly from the video player in client side and presented a data-driven approach to model the metric interdependencies and their complex relationships to user engagement, aiming at developing a predictive model of user QoE in viewing video. This model could guide CDN and bitrates selection to achieve improvement in overall user engagement. The case studies discussed above demonstrate that the next stage of the development of video delivery network highly relys on multidimensional data-driven information plane and correlations learned from data can be utilized to effectively improve network resource allocation on demand, enhance user experience, and reduce cost.

5. ACKNOWLEDGEMENT

This work was supported in part by the National Key Research and Development Program under Grant no. 2016YFB1000102, in part by the National Natural Science Foundation of China under Grant nos. 61672318, in part by the projects of Tsinghua National Laboratory for Information Science and Technology (TNList), in part by the European Seventh Framework Programme (FP7) under Grant no. PIRSES-GA-2012-318939, in part by Natural Science Foundation of China no. 61402343 ad in part by Natural Science Foundation of Jiangsu/Suzhou no. BK20160385.

6. CONCLUSION

The video delivery network in new era is a hot and active research area. In this article, we review the evolution of Internet video ecosystem and summarize the key challenges video delivery network are faced with. There are many efforts invested in this field, including the cooperation between stakeholders, optimization algorithms for video delivery quality, the introduction of cloud computing and the invention of ICN architecture to partially address scalability, QoS and flexibility challenges. In order to fundamentally solve these challenges, we highlight the necessity of data-driven multidimensional information plane for video delivery network. Along with the case studies, we demonstrate significant practical benefits that video delivery network can obtain by using information plane.

7. REFERENCES

[1] Cisco. Cisco global cloud index: Forecast and methodology, 2015-2020. *CISCO White paper*, pages 1–28, 2016.

[2] F. Dobrian, A. Awan, D. Joseph, A. Ganjam, J. Zhan, V. Sekar, I. Stoica, and H. Zhang. Understanding the impact of video quality on user engagement. *ACM SIGCOMM CCR*, 41(4), 2013.

[3] X. Liu, F. Dobrian, H. Milner, J. Jiang, V. Sekar, I. Stoica, and H. Zhang. A case for a coordinated internet video control plane. In *Proc. of SIGCOMM*, pages 359–370, 2012.

[4] Z. Wang, L. Sun, X. Chen, W. Zhu, J. Liu, M. Chen, and S. Yang. Propagation-based social-aware replication for social video contents. In *Proc. of Multimedia*, pages 29–38, 2012.

[5] Pantelis A Frangoudis, Louiza Yala, Adlen Ksentini, and Tarik Taleb. An architecture for on-demand service deployment over a telco cdn. In *Communications (ICC), 2016 IEEE International Conference on*, pages 1–6. IEEE, 2016.

[6] Larry Peterson, Bruce Davie, and Ray van Brandenburg. Framework for content distribution network interconnection (cdni). Technical report, 2014.

[7] J. Jiang, V. Sekar, and H. Zhang. Improving fairness, efficiency, and stability in http-based adaptive video streaming with festive. In *Proc. of CoNext*, pages 97–108, 2012.

[8] T. Huang, R. Johari, N. Mckeown, M. Trunnell, and M. Watson. A buffer-based approach to rate adaptation (evidence from a large video streaming service). In *Proc. of SIGCOMM*, pages 187–198, 2014.

[9] X. Yin, A. Jindal, V. Sekar, and B. Sinopoli. A control-theoretic approach for dynamic adaptive video streaming over http. In *Proc. of SIGCOMM*, pages 325–338, 2015.

[10] S. Akhshabi, L. Anantakrishnan, A. Begen, and C. Dovrolis. What happens when http adaptive streaming players compete for bandwidth? In *Proc. of NOSSDAV*, pages 9–14, 2012.

[11] S. Akhshabi, L. Anantakrishnan, C. Dovrolis, and A. Begen. Server-based traffic shaping for stabilizing oscillating adaptive streaming players. In *Proc. of NOSSDAV*, pages 19–24, 2013.

[12] L. De Cicco, S. Mascolo, and V. Palmisano. Feedback control for adaptive live video streaming. In *Proc. of Multimedia Systems*, pages 145–156, 2011.

[13] V. Adhikari, S. Jain, and Z. Zhang. Where do you "tube"? uncovering youtube server selection strategy. In *Proc. of ICCCN*, pages 1–6, 2011.

[14] V. Adhikari, Y. Guo, F. Hao, M. Varvello, V. Hilt, M. Steiner, and Z. Zhang. Unreeling netflix: Understanding and improving multi-cdn movie delivery. In *Proc. of INFOCOM*, pages 1620–1628, 2012.

[15] H. Liu, Y. Wang, Y. Yang, H. Wang, and C. Tian. Optimizing cost and performance for content multihoming. In *Proc. of SICCOMM*, pages 371–382, 2012.

[16] A. Ganjam, J. Zhan, X. Liu, F. Siddiqi, J. Jiang, V. Sekar, I. Stoica, and H. Zhang. C3: Internet-scale control plane for video quality optimization. In *Proc. of NSDI*, 2015.

[17] Y. Wu, C. Wu, B. Li, X. Qiu, and F. Lau. Cloudmedia: When cloud on demand meets video on demand. In *Proc. of ICDCS*, pages 268–277, 2011.

[18] J. He, D. Wu, Y. Zeng, X. Hei, and Y. Wen. Toward optimal deployment of cloud-assisted video distribution services. *IEEE Transactions on Circuits and Systems For Video Technology*, 23(10):1717–1728, 2013.

[19] Y. Wu, C. Wu, B. Li, L. Zhang, Z. Li, and F. Lau. Scaling social media applications into geo-distributed clouds. In *Proc. of INFOCOM*, pages 684–692, 2012.

[20] Y. T. Yu, F. Bronzino, R. Fan, C. Westphal, and M. Gerla. Congestion-aware edge caching for adaptive video streaming in information-centric networks. In *Proc. of CCNC*, pages 588–596, 2015.

[21] Y. Sun, S. Fayaz, Y. Guo, V. Sekar, Y. Jin, M. Kaafar, and S. Uhlig. Trace-driven analysis of icn caching algorithms on video-on-demand workloads. In *Proc. of CoNext*, pages 363–376, 2014.

[22] T. E. Ng and H. Zhang. Global network positioning: a new approach to network distance prediction. *Computer Communication Review*, 32(1):61, 2002.

[23] Harsha V Madhyastha, Ethan Katz-Bassett, Thomas E Anderson, Arvind Krishnamurthy, and Arun Venkataramani. iplane nano: Path prediction for peer-to-peer applications. In *NSDI*, volume 9, pages 137–152, 2009.

[24] H. Yin, X. Zhang, T. Zhan, Y. Zhang, G. Min, and D. O. Wu. Netclust: A framework for scalable and pareto-optimal media server placement. *Transactions on Multimedia*, 15(8):2114–2124, 2013.

[25] Matthew Luckie, Bradley Huffaker, Amogh Dhamdhere, Vasileios Giotsas, et al. As relationships, customer cones, and validation. In *Proceedings of the 2013 conference on Internet measurement conference*, pages 243–256. ACM, 2013.

[26] Vasileios Giotsas, Matthew Luckie, Bradley Huffaker, et al. Inferring complex as relationships. In *Proceedings of the 2014 Conference on Internet Measurement Conference*, pages 23–30. ACM, 2014.

[27] Syed Hasan, Sergey Gorinsky, Constantine Dovrolis, and Ramesh K Sitaraman. Trade-offs in optimizing the cache deployments of cdns. In *IEEE*

INFOCOM 2014-IEEE Conference on Computer Communications, pages 460–468. IEEE, 2014.

[28] A. Balachandran, V. Sekar, A. Akella, Seshan S., I. Stoica, and H. Zhang. Developing a predictive model of quality of experience for internet video. *ACM SIGCOMM CCR*, 43(4):339–350, 2013.