Guest Editorial
Special Section on Cloud-Based Mobile Media: Infrastructure, Services, and Applications

THE proliferation of smart phones and tablets has fueled a new wave of demand for mobile services. Today's users not only expect ubiquitous network connections, but also seek a rich media experience wherever they go and whenever they need. The thrive of mobile media gives rise to a multifold of technical challenges. It is hard to sustain satisfactory quality of experience over inherently unreliable wireless links. Large-scale, simultaneous delivery of media contents tends to swamp existing mobile network infrastructure. Moreover, a fundamental gap exists between the resource demands of computationally expensive multimedia applications and the capabilities afforded by battery-operated mobile devices.

Recent advances in cloud computing help to bridge this gap. By resorting to the cloud as a reservoir for additional computation and storage, mobile media services can scale better with user demand in an elastic and dynamic manner, at relatively low costs. This new paradigm of *cloud-based mobile media* introduces various issues: cloud resource provisioning and scheduling, media content management, quality-of-service (QoS), user mobility, security, etc. Meanwhile, it opens new avenues. Mobile applications may offload computationally expensive tasks to the cloud. Service providers may pool together statistics of mobile devices at the same region, and leverage such knowledge to improve their media transport and caching strategies.

It is the aim of this special section to report on the latest research that explores various aspects of cloud-based mobile media. Through an open call for papers, we received 40 submissions. Fourteen (14) papers were accepted for final publication after two rounds of highly competitive reviews. The final papers were selected on the basis of originality and significance of the technical work, as well as relevance to the theme topic.

The papers in this special section span a wide range of novel algorithms, techniques, applications, and system-level solutions. They naturally fall into two categories: addressing existing challenges or exploring emerging opportunities.

The topics covered by the first group of six (6) papers are representative of issues faced by cloud-based media systems. In "*Efficient Resource Provisioning and Rate Selection for Stream Mining in a Community Cloud*", Ren and van der Schaar present a mathematical framework for jointly provisioning cloud computing resources and selecting media transmission rates. Their formulation strives to strike the optimal balance between energy consumption, classification performance, and application responsiveness. The paper presents an online learning algorithm that does not require knowledge of wireless channel statistics. Its effectiveness is validated via extensive simulations.

Digital Object Identifier 10.1109/TMM.2013.2261614

The second paper by Zhou and Wang, "*Toward Blind Scheduling in Mobile Media Cloud: Fairness, Simplicity, and Asymptotic Optimality*", studies cloud resource scheduling. The authors consider the blind scenario, with unknown system parameters such as user demand rate and server service time. By formulating the resource scheduling problem as a finite-horizon probability-constrained stochastic optimization, they develop a blind scheduling algorithm that performs well in terms of fairness, simplicity, and asymptotic optimality.

The next three papers shift their attention from managing cloud resources to improving QoS. In "*A Network and Device Aware QoS Approach for Cloud Mobile Streaming*", Lai *et al.* present a candidate solution. A dynamic network estimation module attempts to forecast future bandwidth changes based on past observations. A network and device-aware Bayesian prediction module infers the appropriate video format from features and available resources of mobile devices. A prototype of the proposed system realizes real-time transcoding operations within the map-reduce framework in cloud computing.

A different approach is investigated in "*Design QoS-Aware Multi-Path Provisioning Strategies for Efficient Cloud-Assisted SVC Video Streaming to Heterogeneous Clients*". In this work, Zhu, Li, and Chen present a video delivery framework that integrates scalable video coding (SVC) with multi-path routing. Their design leverages the storage and computing power of the cloud to collect network status and calculate video streaming routes in a centralized manner. Several multipath provisioning algorithms are evaluated for cloud-assisted SVC streaming to heterogeneous clients.

In "*On the Investigation of Cloud-Based Mobile Media Environments with Service-Populating and QoS-aware Mechanisms*", Sardis *et al.* raise the concern that the Internet may not respond fast enough to real-time services. They envision an alternative solution: serving individual geographical regions from small, local clouds. Such a design relieves the Internet from backbone congestion. Meanwhile, users can still access services from other regions, if needed.

The last paper in this group addresses the important issue of security. Wu, Wei, and Deng present a solution for cloud-based content sharing in "*Attribute-Based Access to Scalable Media in Cloud-Assisted Content Sharing Networks*". They build upon the technique of ciphertext policy attribute-based encryption (CP-ABE), and extend the scheme to accommodate multi-level access to scalable media. The proposed scheme is robust against multi-user collusion in the cloud. It supports fast and efficient decryption of mobile devices by offloading heavy-lifting computations to cloud servers.

The second group of eight (8) papers explores opportunities offered by cloud-based mobile media. They either present novel

techniques for existing systems, or propose new applications and services. In "*Optimizing Cloud Resources for Delivering IPTV Services through Virtualization*," Vaneet *et al.* describe how server virtualization technology helps to save peak-time resources in IPTV deployments. The key idea is to serve two types of typical requests—instant channel change (ICC) with stringent latency requirements and Video-on-Demand (VoD) with relatively relaxed delivery deadlines—from the same resource pool. By intelligently shifting VoD requests in anticipation of periodic ICC bursts, the authors demonstrate via trace-driven simulations that their proposed scheme can reduce overall server costs by 24%.

For cloud-based file downloading systems, Zhou, Fu, and Chiu study two operational modes in "*An Adaptive Cloud Downloading Service*". Their analysis reveals that the server mode scales well with large varieties of content requests. On the other hand, the helper mode is more efficient for serving highly concentrated requests. The paper further presents an automatic mode switching algorithm based on these insights.

Chen *et al.* aim at improving mobile video streaming with cloud computing in "*AMES-Cloud: A Framework of Adaptive Mobile Video Streaming and Efficient Social Video Sharing in the Clouds*". Their solution deploys video agents in the cloud on behalf of individual mobile users, and elastically allocates cloud resources for accelerating video encoding and delivery processes. Knowledge of users' social relationships is used to improve content pre-fetching efficiency.

In "*CloudMoV: Cloud-Based Mobile Social TV*", Wu *et al.* utilize the cloud to emulate the living-room video watching experience for a group of disparate mobile users. The proposed platform combines several elements: flexible video transcoding, maximizing battery efficiency of mobile devices, and supporting spontaneous social interactions. CloudMoV is implemented on Amazon EC2 and Google App Engine. Its performance has been evaluated in real-world experiments.

The growing amount of images and videos captured by smart phones has increased the appeal of automatic or semiautomatic management of personal media contents. Large scale image annotation is essential in fulfilling this task. Tao *et al.* study cloud-based image annotation in "*Hessian Regularized Support Vector Machines for Mobile Image Annotation on the Cloud*". Their proposed scheme compresses images using Hamming compressed sensing and conducts semantic annotation through a novel Hessian regularized support vector machine in the cloud. Evaluations of this approach yield promising results on the PASCAL VOC07 dataset.

In "*Cloud-Based Image Coding for Mobile Devices—Toward Thousands to One Compression*", Xue *et al.* exploit the availability of highly correlated contents in the cloud for image coding. They present an image coding scheme that generates a baseline from a down-sampled version of the original image, and encode the scale-invariant feature transform (SIFT) vectors in a predictive manner. The paper reports significantly higher compression ratios over conventional image coding, such as JPEG, JPEG2000, and intra-coding in HEVC, for data sets containing highly correlating images.

In "*GPS/HPS- and Wi-Fi Fingerprint-Based Location Recognition for Check-In Applications over Smartphones in Cloud-Based LBSs*", Bisio *et al.* propose a new Location Recognition algorithm for Automatic Check-In applications (LRACI). LRACI is designed for scenarios where traditional approaches, such as Global and Hybrid Positioning Systems (GPS/HPS), tend to fail. It opportunistically exploits the presence of Wi-Fi access points to complement existing GPS/HPS signals. The design hinges on two key ideas: sliding-window-based filtering of GPS/HPS signals and a novel method for calculating Wi-Fi fingerprints. Its Android-based implementation can achieve a location recognition accuracy of 90%.

The last article, "*Adaptive Mobile Cloud Computing to Enable Rich Mobile Multimedia Applications*" combines an extensive overview of cloud-based mobile media with original technical contributions. In this paper, Wang and Dey sketch out early marketing trends and three major challenges: guaranteeing user experience, managing communication and computation costs, and scaling services to millions of users. Using cloud mobile gaming as a concrete application example, the authors demonstrate the effectiveness of a novel joint rendering and encoder bit rate adaptation (JREA) technique in addressing all three challenges. The paper continues to discuss promising future research directions for fully reaping the benefits of cloud computing in mobile media networks.

The guest editorial team would like to acknowledge all authors for submitting their work to this special section. We appreciate the valuable inputs from our reviewers. Finally, we are grateful to Ms. Rebecca Wollman for her help in coordinating the publication process, and to Prof. Mihaela van der Schaar for her gracious support as Editor-in-Chief.

We hope you will enjoy reading this special section.

CHANG WEN CHEN, *Lead Guest Editor*
University at Buffalo
Dept. of Computer Science and Engineering
Buffalo, NY USA
chencw@buffalo.edu

XIAOQING ZHU, *Guest Editor*
Cisco Systems Inc.
Advanced Architecture and Research
San Jose, CA USA
zhuxq@alumni.stanford.edu

YONGGANG WEN, *Guest Editor*
Nanyang Technological University
School of Computer Engineering
Singapore
ygwen@ntu.edu.sg

JOEL J. P. C. RODRIGUES, *Guest Editor*
University of Beira Interior
Instituto de Telecomunicações
Covilhá, Portugal
joeljr@ieee.org