EMERGING 2017

The Ninth International Conference on Emerging Networks and Systems Intelligence

November 12 - 16, 2017

Barcelona, Spain

EMERGING 2017 Editors

Mark Leeson, University of Warwick, United Kingdom

Ioannis Moscholios, University of Peloponnese, Greece

Pascal Lorenz, University of Haute-Alsace, France

EMERGING 2017

Forward

The Ninth International Conference on Emerging Networks and Systems Intelligence (EMERGING 2017), held between November 12 - 16, 2017, in Barcelona, Spain, constituted a stage to present and evaluate the advances in emerging solutions for next-generation architectures, devices, and communications protocols. Particular focus was aimed at optimization, quality, discovery, protection, and user profile requirements supported by special approaches such as network coding, configurable protocols, context-aware optimization, ambient systems, anomaly discovery, and adaptive mechanisms.

Next-generation large distributed networks and systems require substantial reconsideration of exiting 'de facto' approaches and mechanisms to sustain an increasing demand on speed, scale, bandwidth, topology and flow changes, user complex behavior, security threats, and service and user ubiquity. As a result, growing research and industrial forces are focusing on new approaches for advanced communications considering new devices and protocols, advanced discovery mechanisms, and programmability techniques to express, measure and control the service quality, security, environmental and user requirements.

The conference had the following tracks:
- Technology and networking trends
- Applications and services

We take here the opportunity to warmly thank all the members of the EMERGING 2017 technical program committee, as well as all the reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and effort to contribute to EMERGING 2017. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

We also gratefully thank the members of the EMERGING 2017 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope that EMERGING 2017 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the field of emerging networks and systems intelligence.

We also hope that Barcelona, Spain, provided a pleasant environment during the conference and everyone saved some time to enjoy the unique charm of the city.

EMERGING 2017 Chairs

EMERGING Steering Committee

Carl James Debono, University of Malta, Malta
Raj Jain, Washington University in St. Louis, USA
Ioannis Moscholios, University of Peloponnese, Greece
Henrik Karstoft, Aarhus University, Denmark
Samuel Fosso Wamba, Toulouse Business School, France
Masakazu Soshi, Hiroshima City University, Japan
Robert Bestak, Czech Technical University in Prague, Czech Republic
Dong Ho Cho, KAIST, Republic of Korea
Tamás Matuszka, INDE R&D, Budapest, Hungary
Emilio Insfran, Universitat Politecnica de Valencia, Spain
Mark Leeson, University of Warwick, UK

EMERGING Industry/Research Advisory Committee

Euthimios Panagos, Vencore Labs, USA
Dimitris Kanellopoulos, University of Patras, Greece
Linda R. Elliott, Army Research Laboratory, USA
Jason Zurawski, Lawrence Berkeley National Laboratory / Energy Sciences Network, USA
Rawa Adla, University of Detroit Mercy | Ford Motor Company, USA
Arun Das, Visa Inc., USA
Shaohan Hu, IBM Research, USA
Eric MSP Veith, TU Bergakademie Freiberg, Germany
Xiaochuan Fan, HERE North America LLC, USA

EMERGING 2017
Committee

EMERGING Steering Committee

Carl James Debono, University of Malta, Malta
Raj Jain, Washington University in St. Louis, USA
Ioannis Moscholios, University of Peloponnese, Greece
Henrik Karstoft, Aarhus University, Denmark
Samuel Fosso Wamba, Toulouse Business School, France
Masakazu Soshi, Hiroshima City University, Japan
Robert Bestak, Czech Technical University in Prague, Czech Republic
Dong Ho Cho, KAIST, Republic of Korea
Tamás Matuszka, INDE R&D, Budapest, Hungary
Emilio Insfran, Universitat Politecnica de Valencia, Spain
Mark Leeson, University of Warwick, UK

EMERGING Industry/Research Advisory Committee

Euthimios Panagos, Vencore Labs, USA
Dimitris Kanellopoulos, University of Patras, Greece
Linda R. Elliott, Army Research Laboratory, USA
Jason Zurawski, Lawrence Berkeley National Laboratory / Energy Sciences Network, USA
Rawa Adla, University of Detroit Mercy | Ford Motor Company, USA
Arun Das, Visa Inc., USA
Shaohan Hu, IBM Research, USA
Eric MSP Veith, TU Bergakademie Freiberg, Germany
Xiaochuan Fan, HERE North America LLC, USA

EMERGING 2017 Technical Program Committee

Andrea F. Abate, University of Salerno, Italy Vision and image related trends
Mohd Helmy Abd Wahab, Universiti Tun Hussein Onn Malaysia, Malaysia
Manal Abdullah, King Abdulaziz University, KAU Jeddah, Saudi Arabia
Rawa Adla, University of Detroit Mercy | Ford Motor Company, USA
Smriti Agrawal, Chaitanya Bharathi Institute of Technology, Hyderabad, India
Firkhan Ali Bin Hamid Ali, Universiti Tun Hussein Onn Malaysia, Malaysia
Adel Al-Jumaily, University of Technology, Sydney, Australia
Tulin Atmaca, Telecom SudParis, France
Mozhgan Azimpourkivi, Florida International University, USA
Dirk Bade, University of Hamburg, Germany
Nik Bessis, Edge Hill University, UK
Robert Bestak, Czech Technical University in Prague, Czech Republic

Kechar Bouabdellah, University of Oran 1 Ahmed Benbella, Algeria

Lars Braubach, Complex Software Systems | Bremen City University, Germany

Chin-Chen Chang, Feng Chia University, Taiwan

Dong Ho Cho, KAIST, Republic of Korea

Gianpiero Costantino, Institute of Informatics and Telematics (IIT) | National Research Council (CNR), Italy

Arun Das, Visa Inc., USA

Carl James Debono, University of Malta, Malta

Ramadan Elaiess, University of Benghazi, Libya

El-Sayed El-Alfy, King Fahd University of Petroleum and Minerals, Saudi Arabia

Linda R. Elliott, Army Research Laboratory, USA

Xiaochuan Fan, HERE North America LLC, USA

David R. Gnimpieba Zanfack, University of Picardie Jules Verne, France

Nuno Gonçalves Rodrigues, Polytechnic Institute of Bragança, Portugal

Chunhui Guo, Illinois Institute of Technology, USA

Noriko Hanakawa, Hannan University, Osaka, Japan

William Hurst, Liverpool John Moores University, UK

Shaohan Hu, IBM Research, USA

Abdellah Idrissi, Mohammed V University, Rabat, Morocco

Sergio Ilarri, University of Zaragoza, Spain

Emilio Insfran, Universitat Politecnica de Valencia, Spain

Raj Jain, Washington University in St. Louis, USA

Tauseef Jamal, PIEAS University, Islamabad, Pakistan

Jin-Hwan Jeong, SK telecom, Republic of Korea

Guohua Jin, Advanced Micro Devices, USA

Georgios Kambourakis, University of the Aegean, Greece

Dimitris Kanellopoulos, University of Patras, Greece

Henrik Karstoft, Aarhus University, Denmark

Jalaluddin Khan, King Saud University, Saudi Arabia / Aligarh Muslim University, India

Hyunbum Kim, University of North Carolina at Wilmington, USA

Aykut Koc, ASELSAN Research Center, Turkey

Binod Kumar, JSPM Jayawant Institute of Computer Applications, Pune, India

Sunil Kumar, San Diego State University, USA

Mark Leeson, University of Warwick, UK

Yanjun Liu, Feng Chia University, Taiwan

Elsa Maria Macias Lopez, University of Las Palmas De Gran Canaria, Spain

Zoubir Mammeri, IRIT - Paul Sabatier University, Toulouse, France

Tamás Matuszka, INDE R&D, Budapest, Hungary

Natarajan Meghanathan, Jackson State University, USA

Maura Mengoni, Politechnic University of Marche, Ancona, Italy

Ivan Mezei, University of Novi Sad, Serbia

Panagiotis Michailidis, University of Macedonia, Greece

Ioannis Moscholios, University of Peloponnese, Greece

Fabio Narducci, University of Salerno, Italy

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

A Tracking Assisted Relaying Scheme for Decentralized Wireless Networks

Baohua Shao, Mark S. Leeson

School of Engineering
University of Warwick
Coventry CV4 7AL, UK
e-mail: b.shao@warwick.ac.uk; mark.leeson@warwick.ac.uk

Abstract— **We present a computationally simple approach for mobile node motion recognition and prediction in decentralized wireless networks. The ability to observe and analyze movement data and infer targeted mobile subscriber motion strategies offers performance gains in Delay Tolerant Networks (DTNs). The method here offers new mechanisms for the Store-Carry and Forward Relaying Scheme to create opportunities for the system to increase its overall performance. The motion data are processed by the Kalman Filter (KF) algorithm, which may be viewed as a series of prediction, tracking and smoothing calculations of the movement of mobile subscribers. The outcomes illustrate that the algorithm performs well on prediction and tracking the motion of mobile users. A Java based routing protocol implementation of the KF algorithm offers a delivery probability of at least 0.7 and a mean hop count of between 2.1 and 3.7 for small and large networks, respectively. For modest node densities, between 60% and 80% of the messages will be delivered but after a substantial delay. Thus, the approach described will be of benefit in data collection as part of the Internet of Things and small sensor networks.**

Keywords— Store-Carry and Forward; Decentralized Wireless Networks; Kalman Filter

I. INTRODUCTION

One of the aspects introduced into networking by the use of wireless links is the potential mobility of the communicating nodes but in a classic infrastructure-based wireless network, a fixed access point provides onward routing into a wired network [1]. In recent years, the idea of mobile ad hoc networking without a pre-defined structure or designated central controller has gained in popularity with the use of multiple network hops providing the necessary communication range [2]. In such a decentralized wireless network, each of the mobile nodes is allowed to determine how to assist other network nodes since stable end-to-end paths rarely exist because of node mobility, node sparsity and node connection or disconnection [3]. The prevailing network conditions produce a Delay Tolerant Network (DTN), in which connectivity is provided between a pair of nodes despite the intermittent connectivity and long delays to provide more network flexibility and resilience [4]. The network uses the limited connectivity to forward segments of the payload, which can thus be randomly forwarded to any neighboring nodes. This has the potential to lead to unmanageable system efficiency, intolerable overall delay and considerable energy wastage. As a result, specialized routing paradigms have been developed for DTNs [5] that utilize the Store-Carry and Forward (SCF) relaying scheme [6] that is further discussed in Section II. There has also been substantial interest in the use of predictive information to assist in the SCF process, which will also be discussed in Section II. The mechanisms used to determine the forwarding in SCF will critically influence the wireless network performance. The key metrics include network efficiency, transmission delay, Quality of Service (QoS), energy efficiency and network load distribution. Wireless networks provide subscriber mobility and flexibility, allowing portable device vendors to implement more advanced features than in wired networks. The SCF relaying scheme utilizes this mobility to achieve relaying but the positional uncertainty of the wireless nodes can lead to uncontrollable DTN performance. Thus, subscriber motion prediction offers the prospect of well-managed and optimized relay routing.

The rest of the paper is organized as follows. In Section II, the context for the work is provided by a short overview of the literature. The methods of node movement prediction are introduced in the Section III and then the simulation methods are contained in Section IV, culminating with the results obtained. The final section presents the conclusions and suggestions for future work.

II. RELATED WORK

To deal with the motion of the nodes within a DTN, it has been common to form routing paths between nodes that are in each other's direct communication range [7]. Thus, the network needs to maintain an end-to-end structure whilst its intermediate structure varies with node movement. This is difficult because the variations in node positions constantly change the underlying communication graph and mean that nodes must quickly adapt to the new configurations. One of the methods for solving this problem is link reversal [8], which models the problem as a directed graph, reversing the link directions when needed as a result of motion induced connection loss. Unfortunately, as shown in [8], the time to produce a stable link for communication grows as the square of the number of nodes in the network, limiting the scalability of such algorithms. As a result, the SCF approach [6] was developed, in which intermediate mobile nodes store messages in their local memories if they do not encounter a suitable relay node. The messages are then carried whilst the nodes move until they find an appropriate node to which they can forward their data towards the destination.

Early DTN protocols such as epidemic routing [9] operated without network information to aid their decisions. The target in such an approach is to spread packets rapidly throughout the network without a node selection criterion (that would need extra information). Packets are copied at all node encounters and persist in the network until they reach their destination or exceed a chosen lifetime. Protocol performance drops with increasing load because of the growing demands for storage space and low probability that useful forwarding nodes will be encountered rapidly. Limiting the number of copies permitted was introduced by protocols such as Spray-and-Wait (SnW) [10]. In this method, once the maximum number of copies is reached, the carrying node keeps the packet until it reaches the destination, storage limits are exceeded or the packet times out. To overcome the limitations of the random approach above, many protocols have been developed that collect network information to select relay nodes enhancing delivery probability despite limited storage and energy resources [5]. A well-known example of a protocol that predicts contacts among DTN nodes is PROPHET [11]. This produces a node metric via the number of meetings between nodes; the link weightings between nodes are increased when they meet along with the weightings of other nodes that they have met. The adoption of this method produces an increased delivery ratio but at the price of an increased average packet delay. The information gleaned from node interactions may also be used to detect what can be described as *social* relationships between the network nodes [12]. These formalize the concept that to be considered part of the same community, nodes should be in frequent, regular and long-lasting contact that will suggest promising forwarding paths. For brevity, the summary above naturally leaves out many variations on the themes presented, so the interested reader is referred to [13] for further details and references.

With particular reference to uncertainty in wireless subscriber movement prediction, it is known that given knowledge of a large population, accuracies approaching 90% can be achieved [14]. However, here we need real-time estimation based on limited information. Sometimes, the DTN in question will have movement restrictions such as that considered by Ahmed and Kanhere [15]. They considered operation where public transport networks or street patterns reduced the range of subscriber movement choices to simplify the prediction work. In general, we need to allow the networks nodes more freedom and the approach taken can be reactive or proactive [16]. In the former, nodes report their location to a central network authority such as a base station. However, in the latter, prediction is used and this has the potential to reduce the inevitable latency whilst waiting for location updates. The uncertainty arises from the mobility model extending into the future based on known mobility history data. The success of a mobility model depends on how well it can learn and predict future node locations based on the available scenario history [16]. User movements are to a large degree predictable [17] so the problem becomes one of designing an efficient location prediction algorithm using past data.

Similarly, the idea of using prior probability and Bayesian inference to properly drive a search process in ad hoc delay tolerant networks has been exploited [18]. This use of a generic computable inference mechanism to increase the performance of DTNs has gained popularity in the last few years, culminating in a recent study employing a weighted feature Bayesian predictor that outperforms a naïve Bayesian approach [19]. However, there is no comprehensive and systematic research study on the entire system to improve the network performance by using rigorous prediction and analysis methods. Although Kalman filtering has been used to update connection probabilities [20], the work in [18] was the first adoption of Bayesian inference, in the context of DTN routing. However, the main focus of the paper is on gradient routing in which the message tends to follow a gradient of increasing utility function values towards the destination. Another paradigm has been employed by Talipov et al. [21], who utilize a hidden Markov model to predict to predict the future location of individuals. The inspiration for the scheme is the same as ours and based on the observations of Gonzalez et al. [22] that human trajectories show a high degree of temporal and spatial regularity, and in social environments individuals move subject to a deterministic schedule with only a few random deviations.

III. PREDICTION METHOD

The movement prediction of mobile terminals comprises a series of estimations of moving targets. This can use techniques from problems in different areas such as tracking flying objects using radar. Tracking is a special case of estimation, as the inference of mobile subscriber movement will be represented as a set of complex state space estimation elements [23], each of which records a certain mobile subscriber's position, instantaneous velocity and instantaneous acceleration (or deceleration). For each particular moment or interval, every individual mobile node has its own state data set indicating its state space information forming a state space identification vector. A series of these vectors record the trajectory of a mobile subscriber or preset mobile user group within the network.

All the mobile nodes have the ability to move freely around the radio frequency coverage area, with their random motion forming a random walk stochastic system. The unknown state of the targeted wireless subscriber (denoted by X) is computable based on the observation or measurement (denoted by Y) of mobile subscriber behavior. Prediction is then possible with inference using historical measurements [24].

These potentially computationally burdensome tasks, such as algorithm computation and historical data storage, need to be performed by each individual subscriber's mobile device, such as smartphones, tablets, e-book readers, portable handsets and laptops. The outcomes need to be propagated wirelessly. Each mobile device will have its own limitations on processing capacity, embedded storage memory and particularly wireless bandwidth. Thus, the computerized algorithms need to be simplified and utilized on a minimized scale, which is within mobile device capabilities including the available wireless link bandwidth. Each mobile node only needs to track and predict nodes that can establish direct bi-directional radio connections between the two adjacent nodes. The prediction information is only exchanged among these neighboring mobile nodes. Thus, for prediction, each

mobile node needs to obtain its neighbor node state information nodes by movement tracking.

Here, established Bayesian statistical methods are used to accomplish the moving object motion prediction operation [25]. According to the overall behavior of mobile subscribers, the nodes will be classified into different categories by utilizing different criteria, for instance, non-maneuvering objects and maneuvering objects. If the objects are maintaining a constant velocity so that they may be classified as non-maneuvering objects, then the system is Linear Quadratic Gaussian (LQG) [24] and may be solved using linear system models. The motion of maneuvering objects is normally more dynamic with different accelerations and the trajectory is non-linear so the solution will be more difficult and perhaps only sub-optimal solutions can be obtained [26].

IV. TRACKING STRATEGIES

The tracking problem is actually to estimate the state of moving mobile subscribers (targets) based on the observation data via statistical algorithms. The state of the targets can thus be seen as in a dynamical system [27] with time independent states, forming an autonomous system. The trajectories of targeted mobile subscribers are normally continuous, but observations are made at fixed time intervals and so are taken in a discrete mode. This mathematical statistics status is called the continuous − discrete filtering mode [23], with the discrete observations forming the state space information input.

The classical Bayesian approach provides us with a method to deduce the further states of observed moving objects. Bayes' theorem [25] implies that the mobile node states can be predicted from the observation data, which is the joint probability of the state of event x and the observation of event y divided by the unconditional probability of the observation of event y, which is the normalization factor.

The movement of a mobile subscriber is a random walk [24] obeying the Markov property [28], so the stochastic motion of each mobile node can be treated as a series of Markov process individually. A first order Markov chain can be used for predicting the state space identification of each mobile subscriber step by step. The recursive Bayesian solution is [28]:

$$p(x^k|y^k) = \frac{p(y_k|x_k)}{p(y_k|y^{k-1})} p(x_k|x_{k-1})p(x^{k-1}|y^{k-1}) \ (1)$$

Leading to a state conditional density:

$$p(x_k|y^k) = \int_{x_{k-1}} p(x^k|y^k)dx_{k-1} \qquad (2)$$

In these equations, the superscripts refer to vectors of all x or y values from one to k or k-1 whereas the subscripts denote single instances of x or y.

A. Simulation Model

The targeted system and observation methods are based on linear system models with quadratic system optimization. The wireless system and observation are subject to Gaussian noise so they obey the basic LQG regulator [24]. Hence, the object tracking and movement prediction problem can be solved by a Kalman Filter (KF)

[24]. Equation (2) is the recursive estimation of the state conditional density function and the term $p(x^{k-1}|y^{k-1})$ gives the prior probability density function. In the Bayesian recursive solution, $p(x_k|y^k)$ is a conditional density of the targeted mobile subscriber state $x_k = (x_{k1}, x_{k2}, ..., x_{kn}) \in \mathbb{R}^n$ at the moment k given all the observed data $y^k = (y_1, y_2, ..., y_k)$ with $y_k = (y_{k1}, y_{k2}, ..., y_{km}) \in \mathbb{R}^m$.

The moving object tracking algorithm with noise is:

$$x_k = \mathrm{f}(x_{k-1}) + v_k \qquad (3)$$

where $\mathrm{f}(x)$ is some function of x and v_k is a vector of Gaussian noise.

In practice, the movement of mobile users cannot remain at a constant velocity or absolute steady state but relatively small perturbations occur that can be regarded as Gaussian noise. Given that only a small portion of wireless users will exhibit high mobility [22], such a model is of some utility.

In the decentralized wireless networks designed to date, each mobile node has to observe the movement of other nearby nodes and try to estimate the state to implement the SCF relaying scheme. Here, this state is restricted to the position and velocity of the mobile subscriber wireless nodes. The observation cannot be ideal, and there is always some noise that enters the system. Generally, the KF algorithm is able to deal with two kinds of noise, namely measurement or sensor noise and transition or process noise [29]. Both types of noise are zero mean Gaussian in nature, and the dynamic and observation models are linear Gaussian. The filtering model presented above obeys the basic LQG regulator as mentioned before, so the filtering equation can be expressed as [30]:

$$x_k = \mathbf{A}x_{k-1} + q_{k-1} \qquad (4)$$

$$y_k = \mathbf{H}x_{k-1} + r_k \qquad (5)$$

where x_k is the hidden state vector and y_k is the observation vector at time k, respectively; $q_{k-1} \sim N(0, Q)$ is the transition noise; $r_k \sim N(0, R)$ is the sensor noise.

The movement of the mobile subscriber is described by two-dimensional Cartesian coordinates, so the hidden state vector has four dimensions $x_k = (x_{k1}, x_{k2}, x_{k3}, x_{k4})$. The first two elements capture the position of the mobile node and the second two represent its corresponding velocity. The observation vector is $y_k = (y_{k1}, y_{k2})$.

The matrices within the dynamic model are:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{Q} = \begin{pmatrix} 0.1 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.1 \end{pmatrix}$$

where Δt is one second in the simulations and $Q(i, j)$ is the transition covariance [14].

The matrices in the observation model are:

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

where $R(i,j)$ is the observation covariance [14].

Here, the KF equations can be described as two steps [30]:

(i) prediction:

$$m_k^- = A_{k-1} m_{k-1} \qquad (6)$$

$$P_k^- = A_{k-1} P_{k-1} A_{k-1}^T + Q_{k-1} \qquad (7)$$

(ii) update:

$$S_k = H.P_k^-.H^T + R \qquad (8)$$

$$K_k = P_k^-.H^T.S_k^{-1} \qquad (9)$$

$$m_k = m_k^- + K_k.\{y_k - H.m_k^-\} \qquad (10)$$

$$P_k = P_k^- - K_k.S_k.K_k^T \qquad (11)$$

In which

y_k is the measurement at the time step k;

$\mathbf{P}_k$ is the covariance of a Kalman/Gaussian filter at the time step k;

$\mathbf{P}_k^-$ is the predicted covariance of a Kalman/Gaussian filter at the time step k just before the measurement y_k;

S_k is the innovation covariance of a Kalman/Gaussian filter at step k;

K_k is the gain matrix of a Kalman/Gaussian filter;

m_k is the mean of a Kalman/Gaussian filter at the time step k;

m_k^- is the predicted mean of a Kalman/Gaussian filter at the time step k just before the measurement y_k.

Before the filtering process starts, both the state vector **initial_state** (which is a column vector) and the state covariance vector **initial_V** have to be initialized thus:

$$\mathbf{initial\_state} = \begin{pmatrix} 10 \\ 10 \\ 0 \\ 0 \end{pmatrix}$$

$$\mathbf{initial\_V} = \begin{pmatrix} 10 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 10 \end{pmatrix}$$

B. Algorithm Simulation

To simulate the scenario studied, the true mobile user locations are generated by MATLAB, producing a stochastic linear dynamical system, which is a type of hidden state [29]. This is because the mobile node states cannot be directly measured by neighboring mobile subscribers and KF algorithms are used for estimation. Figure 1 illustrates the results of simulated KF algorithms using 50 individual states in each time step. These are the true states that simulate the real locations of the mobile subscriber during a continuous period of time, and that are represented by the black squares. The trajectory shown by

the black line linking the black squares is the 'real path' of the motion of a certain mobile node. The blue stars indicate the observed location of the mobile device which simulates the measurements from another neighboring mobile terminal. The red crosses show the KF outcomes, processed by the neighboring mobile smart device with the estimated path represented by the red dotted line.
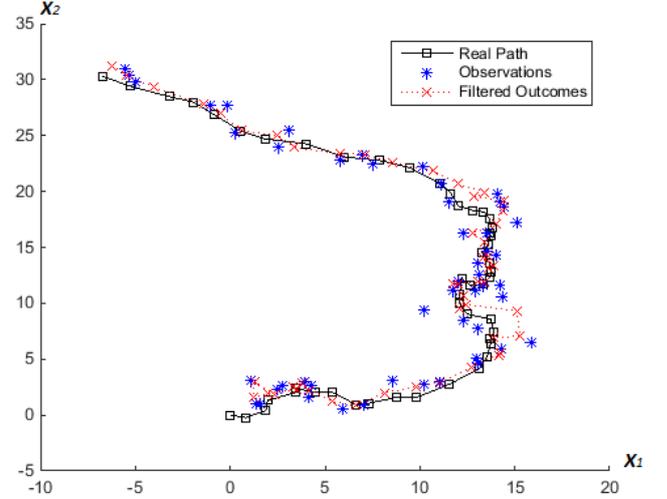
Fig. 1.   Results of the prediction simulation for the filtering model.

It may be seen in Figure 1 that for most of the time, the filtered trace represents the true path well. Only when the mobile user's movement is more dynamic (close to the maneuvering model), particularly the right hand side of Figure 1, does the algorithm have difficulty following the true path. Nevertheless, when the motion of the object exhibits behavior that is close to the non-maneuvering scenario, the outcomes still reflect the real motion of the target very well as in the top and bottom parts of the trajectory, and the mismatched portion is relevant small.

C. Protocol Simulation

The simulation testbed for this part used the Opportunistic Network Environment (ONE) simulator and a JAVA based protocol for the KF routing scheme was developed. For testing the performance, resilience and tolerance of designed protocol, the sample dataset that comes with the ONE simulator package was utilized to simulate a complex wireless network condition, which is the data collected from the downtown Helsinki area. Parameters for the simulation configurations are specified in Table 1. These are chosen to be of the same order as the parameters in [5] with the buffer size large enough that it does not impact performance [5].

TABLE I.   PARAMETERS OF SIMULATION CONFIGURATIONS

Simulation Time (s)	86400
Buffer Size (MB)	50
Packet Lifetime	100 minutes
Message Interval (s)	3, 5, 10, 20, 30, 60
Message Size (kB)	500
Number of Nodes	40, 100, 200, 300, 400, 500

The message interval simulated the information rate of the sender. The parameters for this category tested the

circumstances from a low packet generation rate of 1 packet per minute (67 kbps) to a high packet generation rate of 20 packets per minute (1.33 Mbps). The number of nodes varied the density of the wireless system from a low-density (40 nodes) mobile network to an extremely high-density (500 nodes) system.

In this work, there are four key factors of wireless system that are addressed to evaluate the overall performance of proposed mobile routing strategy, which are: Delivery Probability, Overhead Ratio, Average Latency and Average number of hops.
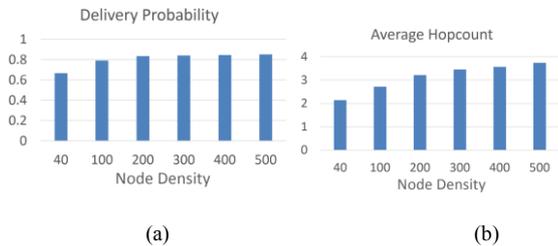

Fig. 2. (a) Delivery probability; (b) average hop count for different network densities.

Figure 2 shows the performance of the proposed protocol at the maximum bit rate considered. It provides good resilience for different network densities and maintains a delivery probability in excess of 0.7 for all circumstances. Moreover, as the algorithm is able to predict the movement of portable nodes, the protocol delivers an average hop count of between 2.1 and 3.7, leading to the involvement of fewer intermediate nodes in the relaying path saving retransmission energy and improving efficiently.

Figure 3 shows that the overhead ratio rises sharply with the number of nodes since there are more possible packet relay candidates. However, there is a corresponding decrease in the average latency as there are more nodes that can complete delivery. The balance of these two factors makes the protocol able to maintain useful performance when the network setup changes.
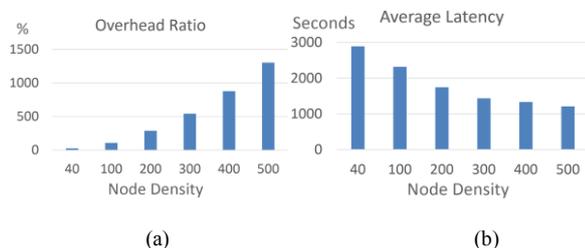

Fig. 3. (a) overhead ratio; (b) average latency for different network densities.

To test the capability of the protocol to deal with various traffic volumes, the packet generation rate in a network comprising 40 nodes was varied. Figure 4 illustrates the variation in delivery probability and hop count as the data rate increases. The former drops with increasing traffic volumes but the KF protocol still maintains a probability of approximately 0.6 whilst the hop count falls from almost three to a little over two with increasing bit rate.
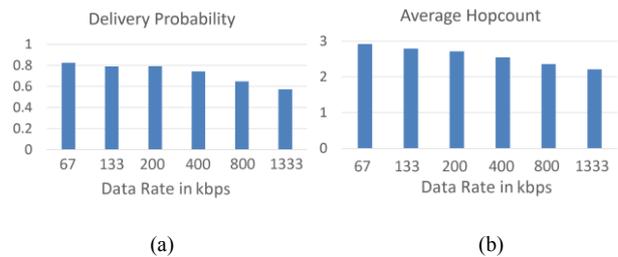

Fig. 4. (a) Delivery probability; (b) average hop count as a funcion of data rate for low node density.

Figure 5 shows that the overhead ratio decreases from 148% to 31% as the bit rate increases but this is accompanied by an increase in average Latency from 1875 seconds to 3153 seconds.
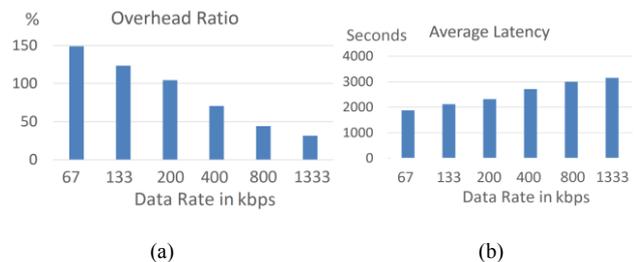
Fig. 5. (a) Overhead ratio; (b) average hop count as a funcion of data rate for low node density.

The KF relaying scheme exhibits a good overall performance which benefits from the portable device movement predication ability allow more packets to arrive successfully at the receiver or be relayed to the correct intermediate nodes. This feature maintains the delivery probability at a high value whilst and keeping the average hop count at a low level.

V. CONCLUSIONS AND FUTURE WORK

The KF is an *optimal recursive data processing algorithm* [24] that provides the online estimation solution to solve the object tracking problem. We have presented a detailed evaluation of the performance of a protocol that utilizes KF algorithm models. This has shown that such an approach enables smart devices to predict and track the motion of a targeted mobile node and assist it to find the next hop as a better or best option for a relaying route. The subsequent routing protocol simulation results proved the theoretical idea. The strengths of the KF are that the algorithm is rather small and simple and thus the majority of mobile devices and sensors are able to process the program. Moreover, the algorithm does not require substantial memory resources to store the movement history of targeted mobile node.

The results indicate that the KF algorithm will face challenges when significant numbers of wireless users fall into the category where a maneuvering model is needed. When the user is moving unsteadily, both the direction and the velocity could be changing at all times; in the case, acceleration or deceleration will be included in future as another dimension of the state vector to indicate the state of the mobile subscriber. However, as the dimension of the inputs becomes high, the calculation volume will substantially increase exponentially. Thus, to let the

algorithm still be available for individual smart devices, the computation time should be taken into account. Nevertheless, the results indicate that at modest node densities, the protocol will deliver between 60% and 80% of the messages but after a substantial delay. Thus, real-time applications will not be well served by the simple approach taken here but it will be useful, for example in data collection as part of the Internet of Things.

Our ongoing work shows that in sparse networks, the KF algorithm exhibits similar delivery probabilities, latency and hops counts to established protocols such as Spray and Wait. Although some of the advantage is lost with dense and complex networks, the protocol's simplicity offers utility to, for example, small sensor networks. Its modest bandwidth requirements also offer advantages in constrained communication environments.

As a classical optimal prediction and tracking algorithm, the KF is suitable for many scenarios, since only small portion of wireless users will exhibit high mobility [17]. The introduction of users who move rapidly according to a random walker model as described by Shang [31] would lead to significant prediction errors. Hence, to broaden the application of this smart relaying scheme to include such very mobile users, other algorithms that can improve the prediction and tracking performance for the maneuvering model, such as the Extended KF (EKF), Unscented KF (UKF), Particle Filter and other potential filtering schemes [28] will be examined in the future.

REFERENCES

[1] J. Loo, S. Khan and A. N. Al-Khwildi, "Mobile Ad Hoc Network" in Mobile Ad Hoc Networks: Current Status and Future Trends, J. Loo, J. Lloret Maur and J. Hamilton Ortiz, Eds. Boco Raton: CRC, pp. 3-17, 2012.

[2] C. M. Conti and S. Giordano, "Mobile Ad Hoc Networking: Milestones, Challenges, and New Research Directions", IEEE Communications Magazine , vol 52, no. 1, pp. 85-96, 2014.

[3] C. S. Jain, K. Fall and R. Patra, "Routing in a delay tolerant network", ACM SIGCOMM Computer Communication Review, vol. 34, no. 4, pp. 145–158, 2004.

[4] A. M. Vegni and T. D. C. Little, "VANETs as an Opportunistic Mobile Social Network", in Opportunistic Mobile Social Networks, J. Wu and Y. Wang, Eds. Boca Raton: CRC, pp. 438-465, 2015.

[5] T. Abdelkader, K. Naik, A. Nayak, N. Goel and V. Srivastava, "A Performance Comparison of Delay-Tolerant Network Routing Protocols", IEEE Network, vol. 30, no. 2, pp. 46 – 53, 2016.

[6] C. C. Sobin, V. Raychoudhury, G. Marfi and A. Singla, "A survey of routing and data dissemination in Delay Tolerant Networks", Journal of Network and Computer Applications, vol. 67,pp. 128-146, May 2016.

[7] I. Chatzigiannakis, "Communication in Ad Hoc Mobile Networks Using Random Walks", Encyclopedia of Algorithms, New York: Springer, pp. 357-363, 2003.

[8] C. Busch and S. Tirthapura, "Analysis of link reversal routing algorithms. SIAM Journal of Computing, vol. 35, no. 2, pp. 305-326, 2005.

[9] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks", Technical Report CS-2000-06, Department of Computer Science, Duke University, 2000.

[10] T. Spyropoulos, K. Psounis and C. S. Raghavendra, "Spray and Wait: An Effcient Routing Scheme for Intermittently Connected Mobile Networks," ACM SIGCOMM Workshop on Delay-Tolerant Networking (WDTN), Aug. 2005, pp. 252-259. ISBN:1-59593-026-4.

[11] A. Lindgren, A. Doria and O. Schelén, "Probabilistic routing in intermittently connected networks", ACM SIGMOBILE Mobile Computing and Communications Review, vol. 7, no. 3, pp. 19-20, 2003.

[12] E. Bulut and B. Szymanski, "Exploiting Friendship Relations for EfficientRouting in Mobile Social Networks", IEEE Transactions on Parallel and Distributed Systems, vol. 23, no. 12, pp. 2254-65, 2012.

[13] P. K. Verma et al., "Machine-to-Machine (M2M) communications: A survey", Journal of Network and Computer Applications, vol. 66, pp 83-105, March 2016.

[14] X. Lu, E. Wetter, N. Bharti, A.J. Tatem and L. Bengtsson, "Approaching the Limit of Predictability in Human Mobility", Scientific Reports 3, art. 2923, 2013.

[15] S. Ahmed and S. S. Kanhere, "Characterization of a Large-scale Delay Tolerant Network," IEEE Conference on Local Computer Networks (LCN), Oct. 2010, pp. 56 – 63.

[16] N. Dimokas, D. Katsaros, P. Bozanis, and Y. Manolopoulos, "Predictive Location Tracking in Cellular and in Ad Hoc Wireless Networks," Mobile Intelligence, L. T. Yang, A. B. Waluyo, J. Ma, J. Tan and B.Srinivasan Eds., New York: Wiley, pp. 163–190, 2010.

[17] I. Burbey and T. L. Martin, "A survey on predicting personal mobility", International J. of Pervasive Computing and Communications, vol. 8, no. 1, pp.5-22, 2012.

[18] S. Ahmed, S.S. Kanhere,"A Bayesian Routing Framework for Delay Tolerant Networks", Wireless Communications and Networking Conference (WCNC), Apr. 2010, pp. 1-6.

[19] K. Jahromi, M. Zignani, S. Gaito and G. P. Rossi, "Predicting encounter and colocation events", Ad Hoc Networks, vol. 62, pp. 11-21, Jul. 2017.

[20] M. Musolesi, S. Hailes, and C. Mascolo. Adaptive routing for intermittently connected mobile ad hoc networks, IEEE WoWMoM, pp. 183–189, Jun. 2005.

[21] E. Talipov, Y. Chon and H. Cha, "Content sharing over smartphone-based delay-tolerant networks", IEEE Transactions on Mobile Computing, vol. 12, no. 3, pp. 581-595, 2013.

[22] M.C. Gonzalez, C.A. Hidalgo, and A.-L. Barabasi, "Understanding Individual Human Mobility Patterns," Nature, vol. 453, pp. 779-782, Jun. 2008.

[23] Y. Bar-Shalom, X. R. Li and T. Kirubarajan, Estimation with applications to tracking and navigation. New York: Wiley, 2001.

[24] P. S. Maybeck, Stochastic Models, Estimation and Control. New York: Academic Press, 1982.

[25] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari and D. B. Rubin, Bayesian data analysis. Boca Raton: Chapman & Hall/CRC, 2014.

[26] N. Ahmed, M. Rutten, T. Bessell, S. S. Kanhere, N. Gordon and S. Jha, "Detection and Tracking Using Particle-Filter-Based Wireless Sensor Networks," IEEE Transactions on Mobile Computing, vol. 9, no. 9, pp. 1332 – 1345, 2010.

[27] E. R. Scheinerman, Invitation to Dynamical Systems. New York: Prentice Hall, 1996.

[28] S. Challa, M. R. Morelande, D. Musicki and R. J. Evans, Fundamentals of Object Tracking. Cambridge, New York: Cambridge University Press, 2011.

[29] S. J. Russell and P. Norvig, Artificial intelligence: a modern approach. 3rd ed. Boston: Pearson, 2010.

[30] S. Särkkä, Bayesian filtering and smoothing. Cambridge: C.ambridge University Press, 2013.

[31] Y Shang, "Consensus in averager-copier-voter networks of moving dynamical agents", Chaos, vol. 27, art. 023116, [9 pages], 2017.

Developing Pedagogical Effectiveness by Assessing the Impact of Simulation Gaming on Education Operation Management:

Experimental Study on Cycles One and Two at Two Private Schools in Lebanon

Nour Issa, Anwar Kawtharani, and Hassan M. Khachfe

School of Education

Business, Educational, & Management Optimization Research Group (BE-MORE)

The Lebanese International University

Beirut, Lebanon

Email: Nour.issa28@hotmail.com, anwar.kawtharani@liu.edu.lb, hassan.khachfe@liu.edu.lb

Abstract— **This study presents a science simulation game, and evaluates in great depth its impact on education operation management. The suggested simulation was empirically examined by investigating the nature and composition of its framework. The descriptive quantitative methodology was appointed to answer the research questions. Data was collected from 308 student participants in both Cycles One and Two, along with their parents, learning support assistants, and teachers. The findings revealed that the students who played the game experienced a deeper level of learning by the action of intrinsic motivation. Furthermore, simple decision-making skills could be acquired with traditional teaching methods, but simulation games were more effective when students had to develop decision-making abilities for managing complex and dynamic situations.**

Keywords-simulations; interactive learning environment; intrinsic motivation.

I. INTRODUCTION

The recent generation of private school students is experiencing the world with their personal computers (PCs). Many have invested much energy playing computer games and are presently highly skilled at learning and applying complex arrangements of standards through game playing. Proserpio and Gioia claimed that the pedagogical method of the new virtual generation varies from the previously used methods. It is significantly more visual, intelligent, and concentrated on critical thinking [1]. While this could be viewed as a danger to the traditional learning methods, it could likewise be viewed as a chance to develop simulation games that empower the learning of management practices and standards.

Creators utilize diverse wordings to characterize simulation technologies that range from top administration, to pilot test programs, to business test systems, to simulation games, to large and small scale universes, to learning research facilities [2]. Simulations and games are broadly utilized in different domains of educational courses. The utilization of simulation and games relies to a great extent on educator's personal capabilities, learning goal and familiarity with the courses instead of their basic knowledge. The confusion between simulations and games has been available from the time that simulation was invented [3]. Since game based learning is becoming more important in teaching, there is a need to verify the utilization of simulation and games.

Despite the fact that there have been numerous endeavors at elucidation, it is still essential to stress the contrasts between the two concepts and to characterize a simulation game. Webster characterizes a simulation as "the illustration of the action or qualities of one framework using another framework, esp. utilizing a PC". strictly speaking; simulation involves the description of a part of reality in view of an improved, and reflective model.

Game is any challenge (play) among players working under imperatives (rules) for a goal, winning, triumph or payoff is a game [4]. Game is along these lines a chance to utilize one's aptitudes and contend with others. The nomenclature additionally proposes a simulation and pleasant action, despite the fact that in academic setting diversions ought not to be utilized chiefly for entertainment. In reality, Abt (1970) alludes to educational games as "significant games." A game is not inevitably a simulation [5].

Simulations, interactive techniques, allow students to practice something that is not quite the same as what they are familiar with [6]. This new experience can prompt more noteworthy appreciation of the content as well. Upon teaching operation management, simulation games embrace a problem-based learning method [7][8]. The problem-based learning method requires educators to introduce students to real-world operation problems that demand to be evaluated and solved.

This paper is organized into five sections as follows: Section two contains a historical view of simulation. Section three portrays a literature review of the pedagogical simulation, through describing the type of simulation used in the study, and effectiveness along intrinsic motivation level. The fourth section investigates the research method upon discussing the data collection, measurements, findings, and analysis of data. The paper concludes with a brief discussion

of the limitation of the research and next steps to be taken into account.

II. HISTORICAL VIEW OF SIMULATIONS

From a long time ago, simulation and games have been employed for training purposes. Their creation is tracked to war games that were used in ancient China. War games, mainly in the sort of wide games, like chess, have been well known. In Germany, during the 17th century, they were changed into more intellectual and composite games [9]. Game and Simulation first appeared in the wide educational spot in the late 1950s. However, Lane declared in 1995 that management diversions and simulations were in the news and their utilization was expanding after fifteen years. Until the mid-1970s, they did not exist in the instructional outline development [10].

In the past, medical models were used as tools in the study of anatomy; mainframe computer-based simulations were generally utilized within the 1960s [11]. Ellington announced in 1981 that computer-based simulations have been used in science education at all levels.

In the last ten years, new simulation games have evolved as teaching tools in marketing, financial management, project management, knowledge management, risk management and microeconomics [12].

III. PEDAGOGICAL SIMULATIONS

A. Types of pedagogical simulations

Computer simulation for pedagogic purposes, which utilizes the computer as a learning domain, is called computer-assisted learning (CAL). Those alleged "training simulations" commonly come in one of three classifications, according to Stančić [13]:
 • "Live" simulation – real individuals utilize the simulated device in the real setting,
 • "Virtual" simulation – real individuals utilize the simulated device in a virtual setting, and
 • "Constructive" simulation – simulated individuals utilize the simulated device in a virtual setting.

It is imperative to notice that in every one of the three cases individuals manage the simulated device and that demonstrates the distinction among simulation and experimentation. Data of science incorporates the greater part of the above-mentioned employments of a computer and classifications of simulation. That is vital for the improvement of long lasting learning, with an exceptionally noteworthy part of between actions. The virtual simulation was employed for this study that is framed around a prototype of reality in which the children act, perform certain roles, and make decisions essential to deal with the intrinsic problematic cases and utilizing specific structured equipment [14][15].

'Helping Plants Grow Well' is the title of simulation employed, which allows students in both grades one and two to investigate various quantities of the variables needed for plant growth. The third grade simulation game is entitled 'Habitat' and teaches students about proper animal habitat, and their position in the food chain. The fourth grade simulation is entitled 'Life Cycle' and illustrates plant parts and their functions. These simulations are present on BBC website [16].

B. Effectiveness along intrinsic motivation

Simulations based reviews have specified that it is a compelling educational technique for theoretical advancement in science and additionally formative appraisal system, and suggested that it would have an advantageous effect on reasonable change. Numerous experimental studies revealed that simulation games intensify their motivation, including ability, intrigue or interest, and endeavors [17-19].

Students who experienced game based learning domain exhibited higher statistical significance of intrinsic motivation than those who acquired education in conventional school environment [20].

Simulations not only help students to comprehend the material better, but also expand student's interest for the content of the material. In 2011, Neumann, Neumann, and Hood [21] examined computer-based simulations in a school statistic's course where 38 students consented to take part. After the class, the students were reached for a 20 minutes telephone call and their responses were put into defined categories when possible. Utilizing the computer based simulation assisted 66 % of students to experience the actual-world application of the specific subject, assisted 60.5 % to comprehend and gain knowledge about the material, created interest and attentiveness in 29 % of the students, motivated 18 % of the students to discover and expanded enjoyment in 16 % of the students.

IV. RESEARCH METHOD

A systematic review comprised of five stages: (1) recognition of research, (2) choice of primary studies, (3) study quality appraisal, (4) data extraction and observing, and (5) data analysis [22] was used. The main research objective was to analyze the impact and execution of simulation games in private schools and to investigate the interrelationships between the barriers and the impact of other contextual factors in the pedagogic environment.

A. Data collection

The target population for the current study was students at cycle's one and two (6 -10 years old) from Family School in Magdouche and National Evangelical School in Nabatieh. Data was collected from March 20, 2017, untill April 25, 2017.

Participants were recruited from the science classes at cycles one and two. The researchers selected the sample for the study using stratified random because Robson [23] contends that sampling theory supports stratified random sampling as an efficient choice since the means of the

stratified samples are likely to be closer to the mean of the population overall.

The method of collecting data included both control/test and pre/post groups, in addition to a hard copy of the questionnaire, which was filled with the assistance of each classroom teacher. A total of 158 responses were collected from control/test group, and 308 responses at pre/post test groups.

B. Measurements

- The control/test group results were compared and discussed to prove the effectiveness of the research. Through obtaining the mean, mode, standard deviation and standard error.
- The student's questionnaire included three-likert scale response (i.e., -1=disagree, 0=neutral, 1=agree). A smiling face indicated that the students agree. A null face indicated that the student is unaware of the answer, and an unhappy face pointed for disagreeing. It was used with all items.
- The questionnaire was divided into five domains, and within lays 21 items. The obtained results present the mean, mode, standard deviation, and coefficient of variation of each item with their total values.
- The pre/post test grades were compared to emphasize on the pedagogic impact of the simulation game. The obtained results presented the mean, standard deviation, and standard error.
- All students had no previous preparation for pre/post test.

C. Findings

The questionnaire was distributed to students, to measure the impact of the simulation game on their academic achievement. Students in cycle one were assisted by the science teachers upon reading the preceding items.

TABLE I.   STUDENTS' PERCEPTION ON THE SIMULATION GAME

Construct	Agree (%)	Neutral (%)	Disagree (%)
Educational goal			
1. The simulation game is helpful and useful for you current lesson.	80.2	11	8.8
2. The simulation game covers the important topics in the lesson.	70.8	21.8	7.5
3. The simulation game has increased your knowledge in the lesson.	80.2	12.3	7.5
4. The simulation game has transferred some practical skills to you.	56.5	20.8	22.7
5. The simulation game is interesting and enjoyable	88.6	8.8	2.6
Knowledge			
6. The simulation game enables you to apply the lesson.	86	8.8	5.2
7. Playing simulation game demands more effort than you expected.	37.8	11.4	50.8
8. Playing simulation game engaged you more in the lesson.	80.8	10.7	8.5
9. The simulation game was good in testing your decision-making.	82.8	9.1	8.1
10. The simulation game provided you with the knowledge that you can use in real life.	81.5	12.3	6.2
Social learning and environment			
11. I worked more with other group members.	69.2	1.9	28.9
12. My group and I dealt with the game challenges perfectly.	62.3	16.3	21.4
13. I had fun while playing the game with my group.	66.2	15.6	18.2
14. Teamwork is important for performing well in the simulation game.	74.6	18.9	6.5
Student perception toward instruction			
15. The simulation game instructions were well organized.	70.8	25.3	3.9
16. The simulation game organization is acceptable.	81.5	14.9	3.6
17. The simulation game was easy to understand and play.	75.3	16.6	8.1
Evaluation			
18. The simulation game results represent your decision.	90.6	6.5	2.9
19. The performance report is easy to read.	55.8	23.4	20.8
20. The time to take the decisions was enough.	57.5	18.8	23.7
21. The animation of the simulation game is helpful.	89.9	6.2	3.9

Table I lists the perceived intrinsic motivation throughout the game. On one hand, the majority of the students (88.6%) agreed that the game developed interest. On the other hand, a few students (2.8%) disagreed with that. Concerning the effort exerted, the students' perception was positive, in which (50.8%) disagreed that the experienced game demanded more effort. Explicitly, (74.6%) agreed that simulation enhances social learning. Most of the students (75.3%) agreed that the game instructions were easy to understand, which made them do better. Only a few (2.9%) of students did not agree that the game results represented their decision, therefore the hypothesis suggesting that simulation games enhance the decision-making capacity of students has been corroborated.

D. Data Analysis

The researchers started with comparing control-test groups at the baseline, through a sample of 158 students from National Evangelical School in Nabtieh. The stratified random sample continued students from both cycle one and two.

TABLE II.   DESCRIPTIVE STATISTICS SUMMARY FOR CCONTROL/TEST GROUP

	Mean	Mode	Standard deviation	Standard error	Sample size
Control group	7.392	7	1.894	0.150	158
Test group	8.512	10	1.534	0.122	158

Table II shows that the average mean of the control group was lower than that of the test group (7.932< 8.512). The average score obtained by the control group is lower than that of the test group (7<10), showing that the group learned through simulation game achieved a better outcome. However, the standard deviation of the control group was higher than that of the test group (1.894> 1.534).

After comparing the control and test groups, which revealed the academic progress of the test group students, the researchers executed a study quality appraisal on pre-post test groups. The same topic was taught to students at both schools, where the pretest was done at the end of the traditional educational method, while the post-test was done at the end of the simulation.

TABLE III. DESCRIPTIVE STATISTICS SUMMARY FOR PRE/POST TESTS

	Mean	Standard deviation	CV	Sample size
Pre-test	7.355	1.986	16.2%	308
Post-test	9.020	1.068	8.7%	308

Table III shows the grade results from both intended schools. The pre and post-test were the same to exclude any possible contamination. The mean value clearly reveals the progress that students made upon the practice of simulation game, in which their average increased by two grades (7.355>9.020). However, the standard deviation was used to measure the amount of variation of the set of pre/post data values. It shows that the variation in pre-test group (1.986) was higher than that of post-test group (1.068). Hence, the simulation game experienced by the students was effective.

At the end of the science course, present students gave their response to the questionnaire.

TABLE IV. DESCRIPTIVE STATISTICS SUMMARY FOR THE STUDENTS' QUESTIONNAIRE

	Mean	Mode	SD	CV
Educational goal	0.654	1	0.494	111.4%
Knowledge	0.580	0.8	0.64	203.26%
Social learning and environment	0.492	1	0.775	168.7%
Students perception toward instruction	0.706	1	0.553	79.5%
Evaluation	0.605	1	0.623	143.6%
Total Whole of the Questionnaire	0.6074	0.96	0.617	141.29%

Table IV summarizes the result of participants' responses to questionnaire items in the five domains: (1) educational goal, (2) knowledge, (3) social learning and environment, (4) student perception toward instruction, and (5) evaluation. This summary indicates that students practice simulation games at their schools with an overall mean of (0.607), mode of (0.96), SD of (0.617), and CV of (141.29%). The highest average value is obtained upon testing the student's perception toward instruction (0.706>0.654>0.605>0.580>0.490), indicating that the practice of simulation games enhance the student's recognition of provided instructions. The use of computerized instructions within the simulation game provided a higher academic achievement [24].

To study the effect of certain variables, the researchers used the ANOVA test; it ia a parametric test used to compare more than two means and to study if the difference is significant or not. For the interpretation, *p*-value is compared with α (error ratio = 5% i.e. 0.05). If *p*-value $> \alpha \rightarrow$ the researchers consider the difference insignificant and vice versa.

Upon performing the ANOVA test, the obtained p-value for students' perspective toward simulation game in relation with learning in a social environment was 0.00002, showing that the null hypothesis is rejected and there is a significant difference. The students who enjoy the simulation game don't necessarily perform better in teamwork.

While answering the hypothesis that tests the relation between simulation game effectiveness and education operation management, the p-value obtained was 0.058, which indicates that the null hypothesis is significant, and therefore, the effectiveness of simulation game is affected by the proper education operation management organization.

V. CONCLUSION

The analyzed results recommended experiencing a simulation game for enhancing operations management in regard to some issues, such as the game engaging quality, appraisal system, and decision-making assessment. Appraisal of group performance should be supplemented with the individual performance assessment to guarantee decency in checking. The utilization of this simulation should be formally coordinated in the operations management course syllabus for consistency over the academic years.

A. Complications and limitations in this study

There were several limitations that the researchers encountered in this study. The two principal limitations are:

1. The lack of studies in Lebanon that have examined simulation, and

2. The insufficiency of prior research studies evaluating the effectiveness of embedding simulations as a part of science education in primary schools.

Moreover, diverse studies discussing dissimilar pedagogical programs and cultures, stated earlier, were homogenous and similar to the findings of this research. A further crucial reason is the lack of any study in Lebanon done at the same time of this research that uses simulation in the classroom to make use of its results and overcome the possible obstacles faced by the researchers.

Consequently, in order to determine the effectiveness of simulation gaming in science education in primary school, further research is needed, which must focus on:

a) Teachers perspective against the utilization of simulation

b)   How encouraging the school administration is in utilizing ICT

c)   Teacher skills in utilizing ICT, especially the ones concerned with simulation gaming.

B. Implications for further study in Lebanon

There is a deficiency of studies in Lebanon and other Arab countries with the close pedagogical environment that evaluates the degree of effectiveness of simulation games in primary schools for science teaching. Consequently, researchers require supplementary studies to decide which factors might reinforce or obstruct the implementation of simulation in science teaching.

ACKNOWLEDGMENT

REFERENCES

[1]   L. Proserpio And A. Gioia, "Teaching the virtual generation. Academy of Management Learning and Education," vol. 6(1), pp. 69–80, 2007.

[2]   E. Clarke, "Learning outcomes from business simulation exercises: Challenges for the implementation of learning technologies," Education + Training, vol. 51 Issue: 5/6, pp.448-459, 2009. Available from: https://doi.org/10.1108/00400910910987246

[3]   M. Lewis and H. Maylor, "Game playing and operations management education," International Journal of Production Economics, vol. 105, pp. 134–149, 2007.

[4]   J. Bloomer, "What have simulations and gaming got to do with programmed learning and educational technology? Programmed Learning & Educational Technology," vol. 10(4), pp. 224–234, 1973.

[5]   C. Abt, "Serious games". New York: Viking Press, 1970.

[6]   A. Struppert, "It's a whole new fun different way to learn," Students' perceptions of learning with an electronic simulation: Selected results from three case studies in an Australian, an American and a Swiss middle school. The International Journal of Learning, vol. 17, pp. 363 – 375, 2010.

[7]   J. Kanet, "Problem-based Learning - Lessons Learned from an Undergraduate Operations Management Program." Paper presented at the POMS 18th annual conference, Dallas, Texas, USA, 2007.

[8]   B. Naik, "Using PBL Assignments in Undergraduate Operations Management Course." Journal of Higher Education THeory and Practice, vol. 11(2), pp. 84-90, 2011.

[9]   J. Wolfe, "A history of business teaching games in English-speaking and post-socialist countries: the origination and diffusion of a management education and development technology," Simulation & Gaming, vol. 24, pp. 446–463, December 1993.

[10] M. Gredler, "Educational games and simulations: A technology in search of a (research) paradigm." In D. H. Jonassen (Ed.), Handbook of research for educational communications and technology, pp. 521-39, 1996.

[11] P. Bradley, "The history of simulation in medical education and possible future directions. Medical Education," vol. 40(3), 2006, pp. 254–262.

[12] H. Gold and S. Gold, Beat the market: an interactive microeconomics simulation. The Journal of Economic Education, vol. 41(2), 2, p. 216, 2010.

[13] H. Stančić, S. Seljan, A. Cetinić, and D. Sanković, "Simulation Models in Education," In Međunarodna Znan. Konf. Futur. Inf. Sci, January 2007, pp. 469-481.

[14] H. Ellington, "Games and simulations – media for the new millennium." In D. Saunders D and N. Smalley. (Eds.), The International simulation and gaming research yearbook. London: Kogan Page Vol. 8, pp. 13-32, 2000.

[15] A. Piu and C. Fregola, "Transcoding Pattern and Simulation Games in Learning Geometry." A Research in Primary School. In S. A. Meijer and R. Smed , (Eds). Frontiers in Gaming Simulation, 2014, pp. 21-28, Switzerland: Springer International Publishing.

[16] "Helping Plants Grow Well" British Broadcasting Corporation (BBC)(http://www.bbc.co.uk/schools/scienceclips/ages/7_8/plants_grow.shtml)

[17] J. Burguillo, "Using game theory and competition-based learning to stimulate student motivation and performance. Computers & Education," vol. 55(2), 2010, pp. 566-575. Retrieved from: http://dx.doi.org/10.1016/j.compedu.2010.02.018.

[18] D. Cordova and M, Lepper, 'Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice'. Journal of educational psychology, vol. 88(4), 1996, p. 715. Retrieved from: http://dx.doi.org/10.1037/0022-0663.88.4.715.

[19] H. Tüzün, M. Yılmaz-Soylu, T. Karakuş, Y. İnal, and G. Kızılkaya, "The effects of computer games on primary school students' achievement and motivation in geography learning. Computers & Education," vol. 52(1), 2009, pp. 68-77. Retrieved from: http://dx.doi.org/10.1016/j.compedu.2008.06.008.

[20] N. Vos, H. Meijden, and E. Denessen, "Effects of constructing versus playing an educational game on student motivation and deep learning strategy use," Computers & Education, vol. 56(1), 2011, pp. 127-137. Retrieved from: http://dx.doi.org/10.1016/j.compedu.2010.08.013

[21] L. Neumann, M. Neumann, and M. Hood, "Evaluating computer-based simulations, multimedia and animations that help integrate blended learning with lectures in first year statistics," Australasian Journal of Educational Technology, vol. 27(2), pp. 274-289, 2011.

[22] B. Kitchenham, "Procedures for Performing Systematic Reviews>" TR/SE-0401, Department of Computer Science, Keele University, 2004.

[23] C. Robson, "Real World Research. A Resource for Social Scientists and Practitioner Researches, 2nd edition. Blackwell: Oxford," 2002.

[24] O. Serin, "The effects of the computer-based instruction on the achievement and problem solving skills of the science and technology students. TOJET: The Turkish Online Journal of Educational Technology," vol. 10(1), January 20111.

Optimization of Power Consumption in SDN Networks

Adrián Flores de la Cruz, Juan Pedro Muñoz-Gea, Pilar Manzanares-Lopez, Josemaria Malgosa-Sanahuja

Department of Information and Communication Technologies

Universidad Politécnica de Cartagena

Campus Muralla del Mar, 30202. Cartagena, Spain

Email: {adrian.flores, juanp.gea, pilar.manzanares, josem.malgosa}@upct.es

Abstract—**Software-Defined Networking (SDN) offers the possibility to carry out a more direct control of network behavior and to interact directly with the elements of the network. In this paper, the problem of optimizing the power consumption in SDN networks is addressed by looking for the most appropriate set of active switches and links, their associated rates, and the number of flow entries at each SDN switch. We present a formulation for this optimization problem and a heuristic algorithm to reduce time complexity in large topologies. Energy saving values of more than 40% are reached using real traffic demands data.**

Keywords–SDN; Energy efficiency; Optimization.

I. INTRODUCTION

Nowadays, the optimization of network power consumption is considered as one promising field of application for Software-Defined Networking (SDN). This new networking technology offers the possibility to carry out a more direct control of network behavior and it gives us the possibility to interact directly with the elements of the network. Considering that strategies, such as turning off network switches, links, or reducing the link rate, can result in energy savings without affecting the service quality, SDN appears as a proper tool to act over network devices and configure them following these prerequisites.

It is known that the highest energy saving is achieved when entire interconnection devices are turned off. However, the energy consumption of data networks can also be minimized by reducing the number of other active elements. For example, this feature can be implemented by putting into a low-power sleep state (sleep mode) elements, such as line cards or port interfaces whenever a link is not tranferring data.

In addition to this, the link rate can also be configured by the controller. Taking into account that a higher link rate means a higher power consumption, ports can be configured with the most appropriate rate considering the traffic rate that is transferred by the link. Finally, it is also necessary to take into account the power consumption associated to flow routing. In this case, this power consumption directly depends on the number of flows configured in the flow table of a switch, as a higher number of flow entries represents a higher number of flows that traverse the corresponding switch.

Different from previous works [1][2], which focus on power-minimization in SDN considering only the number of active links, in this work, we also consider all the previous factors that affect the switch power consumption. That is, in this paper, the problem of optimizing the power consumption in an SDN network is addressed by looking for the most appropriate set of active switches and links, their associated rates, and the number of flow entries at each SDN switch.

The major contributions of this paper are as follows:

- It has been developed an optimization program that minimizes the total switch power consumption considering all the previously presented factors.
- A heuristic algorithm to optimize the network power consumption is also implemented.
- The implemented heuristic is evaluated in different topologies.

The rest of this paper is structured as follows. In Section II, we introduce previous works about the use of SDN technology to reduce the network power consumption. In Section III, we introduce our model and present the formulation of our ILP model. The developed heuristic algorithm is presented in Section IV. Then, the simulation results are analyzed in Section V. Finally, in Section VI we present the conclusions.

II. RELATED WORKS

In [1], the authors faced the problem of optimizing the energy consumption of SDN networks. Such goal is achieved by an energy-aware routing approach that minimizes the number of active links used to route a given traffic demand in SDN. They propose an Integer Linear Program (ILP) model, as well as a heuristic considering the traffic routing requirements. In [2], the previous model is improved. In addition, results showed that the heuristic algorithm converges much faster and it can handle larger network sizes for which the exact model cannot find solutions in reasonable time.

A state-of-the-art study of energy efficiency strategies in SDN is presented in [3]. Some research works, such as [4] and [5] investigate the power consumption of SDN-related system through measurement studies. For instance, the authors of [4] derived power consumption models based on the measurements of two OpenFlow switches, considering the effect of configuration, management, and the managed traffic. In [5], the authors analyze the implications of different software data planes on the power efficiency, as part of Network Function Virtualization (NFV) implementations.

On the other hand, other research works [6][7] explore energy eficiency in SDN-enabled Data Center Networks (DCN). In a typical data center, energy consumption mainly consists of two parts: servers and network devices. In [6] authors investigate how to minimize the energy consumption by carefully scheduling the multi-path routing, according to the data center traffic demands. They take into account the TCAM (Ternary Content-Addressable Memory) size limitation and formulate the problem into an Integer Linear Problem. In [7], the authors study data center energy optimization with joint consideration of Virtual Machine placement and forwarding rule placement. They also formulate the problem in the form of ILP and propose a low-complexity two-phase heuristic algorithm.

III. ENERGY-AWARE APPROACH

According to [8], the power consumption of an OpenFlow switch (P_{switch}) is modeled as:

$$P_{switch} = P_{base} + P_{config} + P_{control} + P_{OF} \quad (1)$$

It consists of the base power P_{base}, the power of the configuration of the switch P_{config} (it depends on the number of active ports and the configured rates), the power consumption of the control traffic $P_{control}$ (it depends on the rate of outgoing *PacketIn* messages -they are a way for the switch to send a captured packet to the controller-, the rate of incoming *FlowMod* messages -they allow the controller to modify the state of an OpenFlow switch-, and their respective energy consumption per packet), and the power consumption of the processed OpenFlow traffic P_{OF} (it depends on the matches and actions of active OpenFlow rules).

A. Network Model

In this paper, we consider a topology as an undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ includes both the host set $\mathcal{H}$ and the set $\mathcal{V}$ of *candidate* SDN-enabled switches, i.e., $\mathcal{N} = \mathcal{H} \cup \mathcal{V}$, and edge set $\mathcal{E}$ represents the set of *candidate* links between the nodes in $\mathcal{N}$.

We also consider a set of offered unicast demands $\mathcal{D}$ between the hosts in $\mathcal{H}$. For each demand d, h_d denotes its known offered traffic.

B. Formulation

We are interested in creating an algorithm that solves the problem that finds (i) the SDN-enabled switches to be activated in the network, (ii) the links to be installed, (iii) their capacities, and (iv) how the traffic must be routed over the links. The optimization target is to minimize the total power consumption of SDN switches in the network.

The problem is formulated as follows:

- Input parameters:
  - $\mathcal{H}$: Set of network hosts.
  - $\mathcal{V}$: Set of *candidate* SDN-enabled switches.
  - $\mathcal{E}$: Set of *candidate* network links. From this information, $\delta^+(v)$ denotes the set of candidate links outgoing from node v, and $\delta^-(v)$ the set of candidate links to v.
  - $\mathcal{D}$: Set of offered unicast demands.
  - $h_d, d \in \mathcal{D}$: Offered traffic of a demand d. From this information, $a(d)$ denotes the origin host of a demand d and $b(d)$ the destination host.
  - P_{base}: The base power of SDN-enabled switches.
  - $P_{config-port}$: The power associated to the activation of both extreme switches ports of a link. It is a term of contributing to the overall P_{config}.
  - $P_{config-speed}$: The power associated with the configured speed rate. It is a term of contributing to the overall P_{config}.
  - $P_{OF-control}$: The power associated to the control and OpenFlow traffic.
  - U: Maximum capacity of a link.
  - M: An auxiliary really big number.
- Decision variables:
  - $z_e, e \in \mathcal{E}$: 1 if candidate link e is actually installed, and 0 otherwise (there is no link there, and then the capacity of this candidate link must be zero).
  - $u_e, e \in \mathcal{E}$: Capacity of candidate link e.
  - $x_{de}, d \in \mathcal{D}, e \in \mathcal{E}$: Traffic of demand d that traverses candidate link e.
  - $\hat{x}_{de}, d \in \mathcal{D}, e \in \mathcal{E}$: 1 if demand d traverses candidate link e, and 0 otherwise.
  - $x_v, v \in \mathcal{V}$: 1 if candidate node v is actually installed, and 0 otherwise.
  - $d_v, v \in \mathcal{V}$: Number of OpenFlow flows installed in candidate switch v.
- Formulation:

$$\min \quad (P_{base}\sum_v x_v + P_{config-port}\sum_e z_e + \quad (2a)$$

$$+ P_{config-speed}\sum_e u_e + P_{OF-control}\sum_v d_v),$$

subject to:

$$\sum_{e\in\delta^+(v)} x_{de} - \sum_{e\in\delta^-(v)} x_{de} = \begin{cases} h_d, & \text{if } v = a(d) \\ -h_d, & \text{if } v = b(d) \\ 0, & \text{otherwise} \end{cases},$$

$$\forall d \in \mathcal{D}, \forall v \in \mathcal{V} \quad (2b)$$

$$\sum_d x_{de} \leq u_e, \quad \forall e \in \mathcal{E} \quad (2c)$$

$$u_e \leq U z_e, \quad \forall e \in \mathcal{E} \quad (2d)$$

$$\sum_{e\in\delta^+(v)} u_e + \sum_{e\in\delta^-(v)} u_e \leq M x_v, \quad \forall v \in \mathcal{V} \quad (2e)$$

$$x_{de} \leq M\hat{x}_{de}, \quad \forall d \in \mathcal{D}, \forall e \in \mathcal{E} \quad (2f)$$

$$\sum_{e\in\delta^+(v)}\sum_d \hat{x}_{de} = d_v, \quad \forall v \in \mathcal{V} \quad (2g)$$

$$u_e \geq 0, x_{de} \geq 0, d_v \geq 0, \quad \forall d \in \mathcal{D}, e \in \mathcal{E}, v \in \mathcal{V} \quad (2h)$$

$$z_e, \hat{x}_{de}, x_v \in \{0,1\} \quad \forall d \in \mathcal{D}, e \in \mathcal{E}, v \in \mathcal{V} \quad (2i)$$

The objective function (2a) minimizes the total power consumption of SDN switches in the network.

Constraints (2b) are the flow conservation constraints. Constraints (2c) mean that for each link, the traffic carried in the link is less or equal than its capacity (and thus, no link is oversubscribed). Constraints (2d) make that (i) if a candidate link e is off ($z_e = 0$), then there cannot be a capacity in it ($u_e = 0$), and (ii) if a candidate link is on ($z_e = 1$), the link capacity is limited to U.

Constraints (2e) make that if a candidate switch v is off ($x_v = 0$), then the capacity of its outgoing and incoming links has to be equal to zero ($u_e = 0$). Constraints (2f) make that if a demand d does not traverse a link e, then the associated traffic of the demand has to be equal to zero ($x_{de} = 0$).

Constraints (2g) set the number of OpenFlow flows installed in candidate switch v. Finally, constraints (2h) forbid that a link carries a negative amount of traffic of a demand, since this has no physical meaning.

Require: $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ network graph, $\mathcal{D}$ data traffic demands
Ensure: $\mathcal{V}$ active switches, $\mathcal{X}$ active links, $\mathcal{U}$ links capacity,
$\quad$ $\mathcal{F}$ flows in each switch
1: **for all** demands $\in \mathcal{D}$ **do**
2: $\quad$ **if** first demand **then**
3: $\quad\quad$ path=SHORTESTPATH(demand)
4: $\quad\quad$ CALCULATEPOWER(path, demand)
5: $\quad\quad$ UPDATE($\mathcal{V}$, $\mathcal{X}$,$\mathcal{U}$,$\mathcal{F}$)
6: $\quad$ **else**
7: $\quad\quad$ $\mathcal{P}$=ALLPOSSIBLEPATHS(demand)
8: $\quad\quad$ **for all** paths $\in \mathcal{P}$ **do**
9: $\quad\quad\quad$ **if** AVAILABLECAPACITY(path) **then**
10: $\quad\quad\quad\quad$ CALCULATEPOWER(path, demand)
11: $\quad\quad\quad$ **end if**
12: $\quad\quad$ **end for**
13: $\quad\quad$ SELECTMINIMUMPOWERPATH($\mathcal{P}$)
14: $\quad\quad$ UPDATE($\mathcal{V}$, $\mathcal{X}$,$\mathcal{U}$,$\mathcal{F}$)
15: $\quad$ **end if**
16: **end for**

Figure 1. Heuristic algorithm

1: **for all** links $\in$ path **do**
2: $\quad$ UPDATEACTIVELINKS(link, demand)
3: $\quad$ UPDATEINSTALLEDFLOWS(demand)
4: $\quad$ **if** origin(link) **or** destination(link) $\notin \mathcal{V}$ **then**
5: $\quad\quad$ UPDATEACTIVESWITCHES( )
6: $\quad$ **end if**
7: $\quad$ UPDATETOTALPOWER(demand)
8: **end for**
9: **return** totalPower

Figure 2. CALCULATEPOWER(path, demand)

This optimization problem has been implemented in a network planning tool that is publicly available through GNU public license, called Net2Plan [9]. The contributed Java Optimization Modeler (JOM) library allows Java to interface from Net2Plan with GLPK and IPOPT optimization engines. In this way, it is possible to create and configure custom optimization problems within the Net2Plan environment.

However, the previous optimization problem becomes challenging to be solved even on medium-scale topologies. This is because the difficulty of this problem is NP-Hard, so the consumption of resources grow exponentially with the network size. To solve this problem in an alternative way, we develop a heuristic algorithm in the next section.

IV. HEURISTIC ALGORITHM

The proposed algorithm is shown in Figure 1. The main loop of the algorithm consists in determining the minimum power path for each traffic demand, that is, the path that suposses a minimum increment in the total power consumption in the network. In the case of the first demand (line 2), the selected path is the shortest one. Then, the power consumption associated to the configuration of this path in the network is calculated, and the sets of active switches ($\mathcal{V}$), active links ($\mathcal{X}$), links capacities ($\mathcal{X}$), and flows in each switch ($\mathcal{F}$) are updated.

For the rest of traffic demands (line 6), first of all, the set of admissible paths associated with each demand is obtained.

Among all these paths with enough capacity to establish the new connection, we select the path that has a lower increment in the total power consumption in the network (line 13). Then, the sets of active switches ($\mathcal{V}$), active links ($\mathcal{X}$), links capacity ($\mathcal{X}$), and flows in each switch ($\mathcal{F}$) are updated.

Figure 2 shows the method to obtain the power increment that supposes the establishment of a specific demand in a certain path. Line 2 considers if a new link must be activated or an existing link needs to increment its configured speed. Line 3 considers the energy consumption required by the adding of new flow entries in the switches that form the path. Line 5 takes into account if it has been necessary to activate new switches.

V. SIMULATIONS AND RESULTS

To evaluate the performance of our algorithm, a simulator in Python language has been developed. Network topologies are obtained using the NetworkX [10] library. NetworkX is a Python software package that is used for the creation and study of the structure and performance of complex networks. Our tests are performed in two different topologies, one with $|\mathcal{V}| = 20$ candidate switches and another with $|\mathcal{V}| = 70$. On the other hand, the traffic demands have been obtained from SNDlib library [11].

Figure 3 shows the total power consumption in a network with 70 nodes, as a function of the number of demands. In this figure, we can observe that the power consumption using the heuristic is lower than the power consumed when all the swiches and links are active. The saved power values are between 25% and 50%. On the other hand, in this figure it is clear that the increment in the number of demands represents an increment in power consumption.

Figure 4 shows the total power consumption in a network with 20 nodes. In this figure, we can observe the same patterns as in Figure 3. In this figure we can see how the heuristic continues to give us better performance regarding scenarios where only the links that are not used are disconnected. When the number of demands is small, the saved power values are similar to those presented with $|\mathcal{V}| = 70$. However, when the number of demands increases, it can be seen that the percentage of saved power is lower in the case $|\mathcal{V}| = 20$. This is due to the fact that in the case of $|\mathcal{V}| = 20$, as demand between nodes increases, most nodes in the network must be activated to supply such demand. While in the case of $|\mathcal{V}| = 70$, having more nodes available, we can try to route the demands in a higher number of available routes, giving us a higher percentage of saved power.

Figure 6 shows the increment in the total power when new demands are activated in the network. It can be seen that when the first demands are established, the increment in the total power is high. This is because it is the necessary to activate several switches to connect the source and the destination nodes. As new demands are established, the switches have already been activated, so the required power is not so high. There are some peaks in the last demands, though. This is because there is a new source or destination node that has not been activated before, or that simply the links that were already connected cannot give the necessary performance to satisfy the demands and it is necessary to look for another route with a new switch that was not connected.
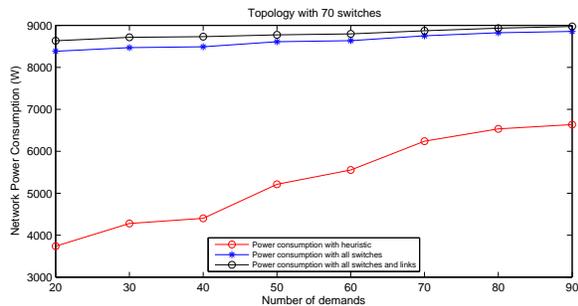
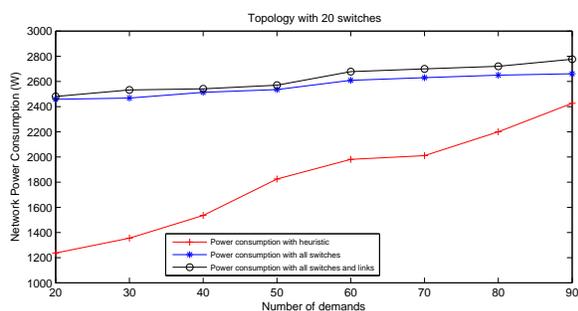Figure 3. Total power consumption in a network with 70 nodes.

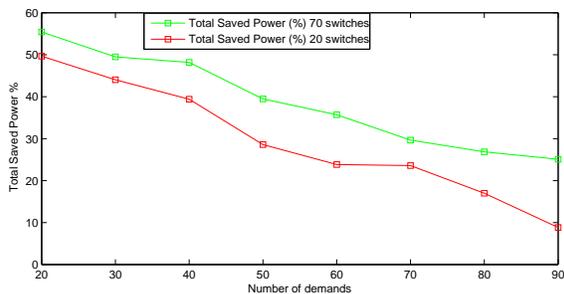Figure 4. Total power consumption in a network with 20 nodes.

Figure 5. Percentage of saved power.

VI. Conclusion

In this paper, we have proposed an approach that minimizes the total power consumption in SDN networks. To achieve this, first of all an optimization program has been formulated. In this program, different constraints have been modeled and implemented. We have also developed a heuristic algorithm that has been evaluated in two different topologies with real traffic demands. Based on experimental simulations, we have proved that our apporach achieves energy savings of up to
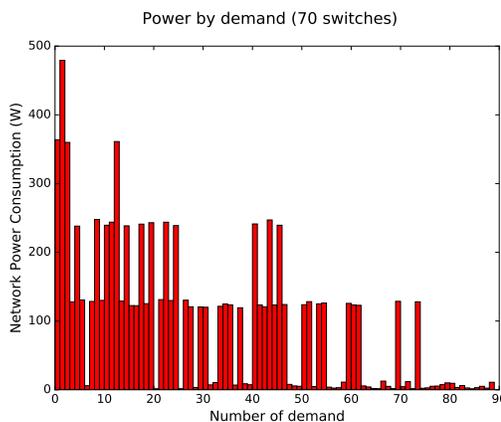
Figure 6. Increment in the total power when new demands are activated.

50%. As future work, we plan to implement this heuristic in Smart Cities environments, where SDN is a key element.

References

[1] A. Fernandez-Fernandez, C. Cervello-Pastor, and L. Ochoa-Aday, "Improved energy-aware routing algorithm in software-defined networks," in *2016 IEEE 41st Conference on Local Computer Networks (LCN)*, Nov 2016, pp. 196–199.

[2] ——, "Achieving energy efficiency: An energy-aware approach in sdn," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–7.

[3] B. G. Assefa and O. Ozkasap, "State-of-the-art energy efficiency approaches in software defined networking," in *SoftNetworking'15*, April 2015, pp. 268–273.

[4] F. Kaup, S. Melnikowitsch, and D. Hausheer, "Measuring and modeling the power consumption of openflow switches," in *10th International Conference on Network and Service Management (CNSM) and Workshop*, Nov 2014, pp. 181–186.

[5] Z. Xu, F. Liu, T. Wang, and H. Xu, "Demystifying the energy efficiency of network function virtualization," in *2016 IEEE/ACM 24th International Symposium on Quality of Service (IWQoS)*, June 2016, pp. 1–10.

[6] D. Zeng, G. Yang, L. Gu, S. Guo, and H. Yao, "Joint optimization on switch activation and flow routing towards energy efficient software defined data center networks," in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.

[7] H. Y. et al.i, "Joint optimization of vm placement and rule placement towards energy efficient software-defined data centers," in *2016 IEEE International Conference on Computer and Information Technology (CIT)*, Dec 2016, pp. 204–209.

[8] F. Kaup, S. Melnikowitsch, and D. Hausheer, "Measuring and modeling the power consumption of openflow switches," in *10th International Conference on Network and Service Management (CNSM) and Workshop*, Nov 2014, pp. 181–186.

[9] P. P.-M. et al., "Net2plan-aire: gained experience in an ad-hoc sdn development for a metro carrier network," in *17th International Network Strategy and Planning Symposium (Networks 2016)p*, 2016, pp. 181–186.

[10] "Networkx," http://networkx.github.io/, accessed: June 2017.

[11] "Sndlib," http://sndlib.zib.de/, accessed: June 2017.

Segment Duration Based
HTTP Adaptive Streaming Scheme for UHD Content

Heekwang Kim and Kwangsue Chung
Department of Electronics and Communications Engineering
Kwangwoon University
Seoul, Korea
Email: hkkim@cclab.kw.ac.kr, kchung@kw.ac.kr

Abstract—**Due to the explosive growth of the mobile devices and development of the network technologies, HyperText Transfer Protocol (HTTP) adaptive streaming has become a new trend in video delivery to provide efficient multimedia streaming services. Another important development is the provision of high quality contents, such as Ultra High Definition (UHD) videos. UHD content supports longer Group Of Pictures (GOP) size, as well as higher spatial resolution and higher frame rates. Due to these characteristics, the UHD content is characterized by the longer segment duration and higher bandwidth requirements in HTTP adaptive streaming system. In an HTTP adaptive streaming system, the long segment duration, which is the characteristic of UHD content, degrades the responsiveness to network fluctuations. It increases the video stalling period and degrades the user Quality of Experience (QoE). In this paper, we propose the segment duration based HTTP Adaptive streaming scheme for UHD contents. In our scheme, we assign different segment duration based on each quality level and propose the rate adaptation scheme to improve the QoE for UHD content. Using the simulation, we prove that our scheme significantly reduces the playback discontinuity.**

Keywords-HTTP adaptive streaming; UHD (Ultra High Definition) content; QoE (Quality of Experience)

I. INTRODUCTION

Due to the development of the network technologies and the explosive growth of the mobile devices such as smart phone and tablet PC, there have been significantly increasing demands for multimedia streaming services. Video traffic is a growing fraction of Internet traffic. Because of video traffic growth, HyperText Transfer Protocol (HTTP) adaptive streaming is receiving attention to provide efficient multimedia streaming services. In the HTTP adaptive streaming, a video content is encoded at multiple bitrates and the encoded video content is segmented into small parts of fixed durations. When the HTTP adaptive streaming player starts video streaming, it requests a segmented video content by sending an HTTP GET message. Each video segment is downloaded over HTTP and TCP using the conventional HTTP web servers. Typical implementations of the HTTP adaptive streaming are Apple HTTP Live Streaming, Microsoft Smooth Streaming, and Adobe dynamic Streaming. Also, Dynamic Adaptive Streaming over HTTP (DASH) is defined in the standard [1].

Another important development is the provision of high quality contents, such as Ultra High Definition (UHD) videos attracting the attention of content providers. The UHD content format is defined in the Recommendation ITU-R BT.2020 and SMPTE ST 2036-1 [2]. Compared to High Definition (HD), UHD content supports longer Group Of Pictures (GOP) size, as well as higher spatial resolution and frame rates to reduce the data size. The UHD content is characterized by the long segment duration and higher bandwidth requirements in the HTTP adaptive streaming system.

In HTTP adaptive streaming system, a client can select the quality level of next segment based on the quality adaptation method to cope with bandwidth fluctuations. The long segment duration, which is the characteristic of UHD content, degrades the responsiveness to network fluctuations [3]. It increases the quality switching delay and video stalling period (also called as buffer freezing). These degrade the user Quality of Experience (QoE).

In this paper, we propose a segment duration based HTTP adaptive streaming scheme for UHD content. The proposed scheme assigns different segment duration according to each quality level and requests the extra segment of the selected quality by the rate adaptation algorithm to improve the responsiveness of UHD content. The proposed rate adaptation algorithm determines the number of extra segments and the quality at which each of these segments should be downloaded. The rest of the paper is organized as follows. The related works are presented in Section II. The proposed scheme is presented in Section III. The experimental results are provided in Section IV, and finally the concluding remarks are given in Section V.

II. RELATED WORKS

Figure 1 illustrates the architecture of HTTP adaptive streaming. An HTTP server consists of three parts, which include the *media segments*, *manifest file*, and *HTTP module*. A multimedia content is encoded with multiple qualities. Each encoded content is divided into small segments of the fixed duration which start with an I-frame and do not reference frames from the surrounding GOPs. The segments are individually addressable by unique Uniform Resource Identifier (URIs). The *manifest file* specifies the content characteristics (e.g., bitrate, codec information, framerate, segment duration, resolution and URI). The *HTTP module* at the server sends the media segments according to request from
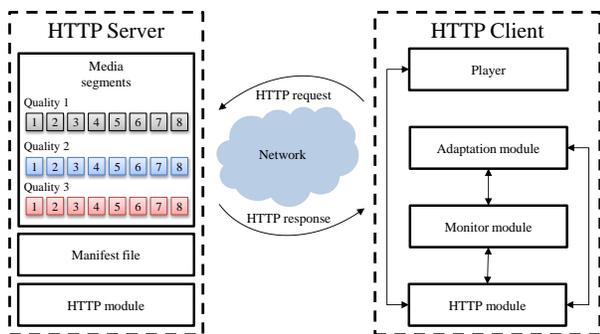
Figure 1. Architecture of HTTP adaptive streaming system.

the HTTP client. An HTTP client consists of four parts, which include the *player*, *adaptation module*, *monitor module*, and *HTTP module*. The *monitor module* estimates the throughput during the download of the segments. The *adaptation module* selects suitable bitrate for adapting the dynamic network conditions depending on the received manifest file and the measured condition, which is the estimated throughput or playback buffer occupancy. The *HTTP module* at the client sends the requests to the HTTP server according to the selected quality. When the HTTP adaptive streaming client starts video streaming, it requests the manifest file using HTTP GET message to the HTTP server. According to this request, the HTTP server sends the manifest file to the client. After receiving the manifest file, the client requests an appropriate segment using throughput measured in monitor module. Then, the HTTP server sends the requested segment. Finally, the *player* plays the received segments.

In order to improve the QoE and Quality of Service (QoS), many adaptation schemes are proposed for HTTP adaptive streaming. The adaptation schemes can be broadly divided into two categories. In the first category, the HTTP adaptive streaming switches the video quality according to the current state of the network, such as throughput [4][5]. Segment throughput is calculated as the ratio of the segment size to the time that it takes to download the segment [6]. The moving average of the throughput of previous segments is used to estimate the throughput [7]. In the second category, buffer occupancy is used in order to provide the seamless playback. The buffer is divided into predefined ranges and different decisions are taken to select the video rates when the buffer level stays in different ranges [8][9]. The method in [9] is more stable as compared to the method in [8] but it is late to react to the changes in the throughput as it waits for the playback buffer to reach a threshold before selecting a higher video rate.

III. SEGMENT DURATION BASED HTTP ADAPTIVE STREAMING SCHEME

In this section, we design the architecture of the proposed HTTP adaptive streaming system for UHD streaming services. Then, we also present a rate adaptation scheme for seamless playback of the UHD content.

A. Architecture of the proposed streaming system

Figure 2 shows the architecture of the proposed streaming system. In the proposed system, the server consists of the *media preparation* module, *content annotation* module and *segment scheduler* module.

- The *media preparation* module provides the tools for encoding and encapsulation so that the content can be presented and delivered efficiently to the client in the pre-defined format. The module encodes the video stream into multiple representations and divides the representation into small segments of different fixed duration according to quality level.
- The *manifest file* module provides metadata file or manifest file which contains the information about the characteristics of the stored multimedia content. The content information contains the bitrate, URI, and segment duration of each representation.
- The *segment scheduler* module is responsible for sending segments using multiple connections according to the request message.
- The *HTTP module* at the server sends the media segment based on requested message by client.

In the proposed system, the client consists of the *media player*, *buffer controller*, *manifest file parser*, *monitoring*, *request scheduler* and *quality adaptation* modules.

- The *player* module provides the tools which play and control the multimedia to the client.
- The *manifest file parser* module analyzes the manifest file received from the server.
- The *buffer controller* module stores the segments in the each buffer according to their connections. Further, this module finds the beginning time of the segments to be played next.
- The *monitor module* measures the throughput while downloading each segment and monitors the buffer occupancy while the client plays the received segments.
- The *adaptation module* selects suitable bitrate to request next based on the received manifest file and the measured context information.
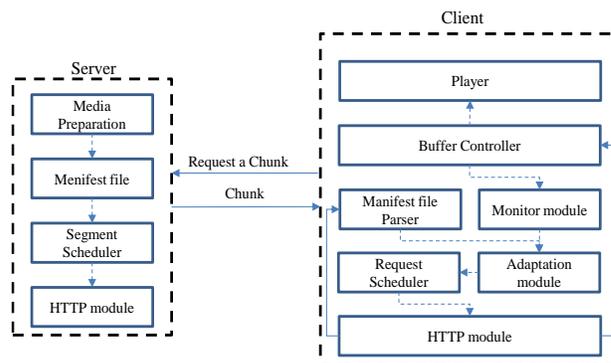

Figure 2. Architecture of the proposed streaming system.

- The *request scheduler* module determines the time to request segment based on the buffer occupancy of the client.
- The *HTTP module* at the client receives the media segment from the server.

B. Segment duration based rate adaptation scheme

We assign the segment duration differently according to each quality level and present the rate adaptation scheme that provides uninterrupted playback even when network bandwidth is decreased.

Figure 3 shows the assigned segment duration for seamless playback of the UHD content. The segment durations are organized hierarchically according to each quality level. The segment duration of the UHD content is 8 seconds due to its longer GOP size. To improve the responsiveness of the network conditions and minimize the wasted bandwidth, the segment duration of other quality level with the exception of UHD content is reduced by half compared to higher quality. For example, the segment durations of Full HD (FHD), HD and Standard Definition (SD) are 4, 2 and 1 seconds respectively. These assignments can quickly increase the quality from SD to UHD by improving the responsiveness of the network conditions.
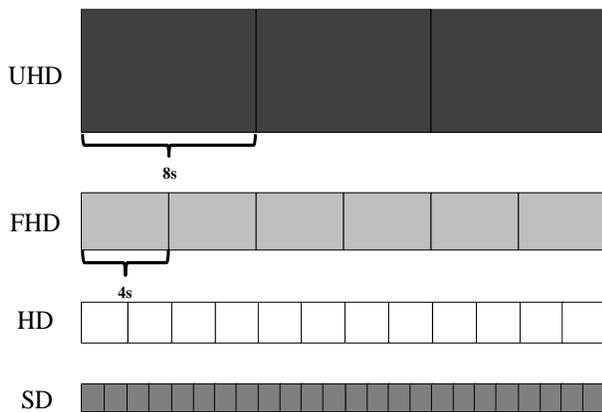

Figure 3. Segment duration assignment.

Figure 4 shows the proposed rate adaptation scheme for seamless playback of the UHD content. The top of the Figure 4 shows the proposed scheme when the bandwidth decreases while downloading the segment. The bottom of the Figure 4 shows the playback scenario of the client. If the requested segment is not downloaded within the segment duration, the client simultaneously requests an extra segment of a lower quality than previously requested using multiple connections while the buffered segments are being played. The extra segment is an additional segment with the same playback time as the segment currently being downloaded. If the lower quality segment gets downloaded before the previously requested segment, the client plays the extra segment while downloading the remaining data of previously requested segment. By playing the extra segment, the client can stream the video smoothly without buffer freezing. After the
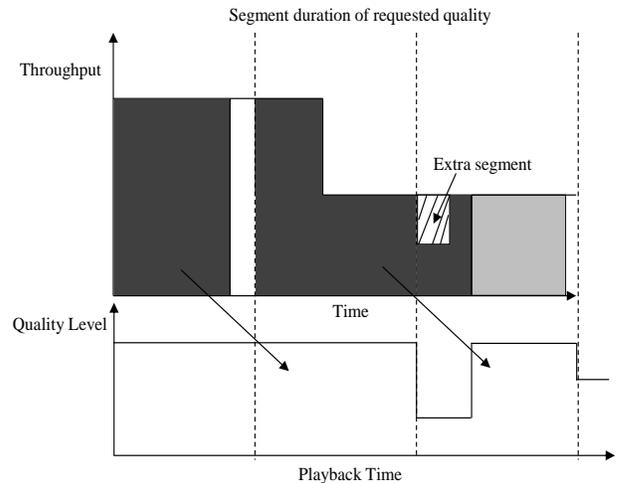

Figure 4. Segment duration based rate adaptation scheme.

remaining data of the higher quality level is downloaded, the proposed system is looking for the time to switches to the higher quality segment.

To improve the user QoE, we present the rate adaptation algorithm for the extra segment. We first determine whether to request an extra segment every segment duration of each quality level. This decision is made by comparing the available buffer level and the required download time. The required download time is calculated as in (1).

$$t_{required} = \frac{S_{remain}}{th_{est}} \tag{1}$$

$t_{required}$ denotes the required download time, S_{remain} denotes the size of remaining data to download, and th_{est} denotes the estimated throughput. The size of remaining data to download is calculated as in (2).

$$S_{remain} = R_{origin} \times t_{origin} - S_{complete} \tag{2}$$

R_{origin} denotes the video rate of original segment, which is the higher quality level, t_{origin} denotes the segment duration of original segment, and $S_{complete}$ denotes the size of downloaded data. The proposed scheme has a condition for decision as in (3).

$$t_{required} > t_{buf} \tag{3}$$

t_{buf} denotes the buffering time as the available download time. If the time required to download the remaining data is larger than the buffering time, the client will experience playback interruptions. Therefore, the client needs to request the extra segment. The proposed scheme also determines the quality level of the extra segment. The quality of extra segment is calculated as in (4).

$$R_{extra} = \max R_k$$
$$, where \; \frac{S_{remain} + n(R_k \times t_k)}{th_{est}} < t_{buf} + n \times t_k, n < \frac{t_{origin}}{t_k} \tag{4}$$

R_{extra} denotes the video rate of extra segment, R_k denotes the video rate of k-th quality level, t_k denotes the segment duration of k-th quality level, and n denotes the number of extra segments. This equation means that the quality adaptation module selects the extra segment that is encoded with the maximum bitrate and avoids buffer underflow.

IV. PERFORMANCE EVALUATION

This section presents the simulation results for the proposed scheme. To evaluate the performance of the proposed scheme, we implement it in the Network Simulatior-3 (ns-3). The simulation topology is shown in Figure 5. In this simulation, the bottleneck link is set at 20Mbps. The segment duration of the conventional HTTP adaptive streaming scheme is 8 seconds for all quality level. The server has 4 different pre-encoded video qualities (430Kbps, 1500Kbps, 2700Kbps, and 10000Kbps). We compare the conventional HTTP adaptive streaming scheme and the proposed scheme. The simulation is run for 400 seconds. In order to generate the cross traffic, we inject 10Mbps of Constant Bitrate (CBR) traffic between the server and client at 200 second.
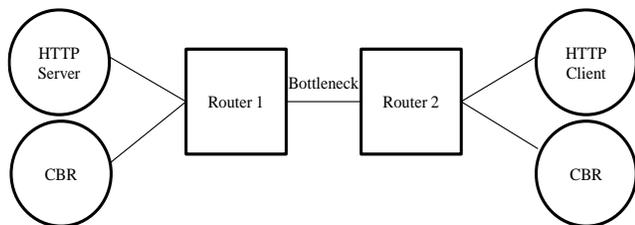
Figure 5. Simulation topology.

Figure 6 shows the playback video level and buffer occupancy of the conventional HTTP adaptive streaming scheme. The conventional HTTP adaptive streaming scheme experiences buffer underflow after generating the cross traffic because the client responds slowly to network changes due to long segment duration. On the other hand, Figure 7 shows the playback video level and buffer occupancy of the proposed scheme. In the proposed scheme, buffer underflow does not occur after generating the cross traffic because the proposed scheme requests an extra segment by predicting buffer underflow. The proposed scheme downloads an extra segment at 207 second and plays the extra segment while downloading the original segment. After receiving the original segment, the player switches to the quality of original segment. Simulation results show that the proposed scheme does not experience buffer underflow and provides a better QoE than conventional HTTP adaptive streaming scheme.

V. CONCLUSION

The UHD content is characterized by the long segment duration and higher bandwidth requirements in the HTTP adaptive streaming system because it supports not only higher frame rates but also longer GOP size. These characteristics degrade the user QoE due to lower responsiveness of the network changes. In this paper, we propose the segment duration based HTTP adaptive streaming scheme for UHD
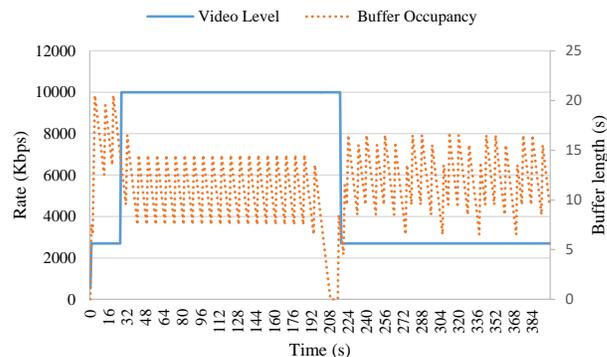
Figure 6. Playback video level and buffer occupancy of the conventional HTTP adaptive streaming scheme.
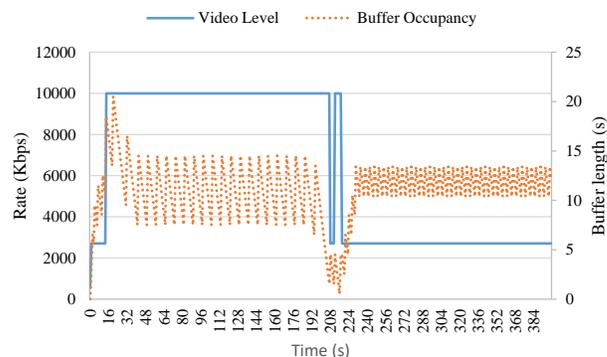
Figure 7. Playback video level and buffer occupancy of the proposed scheme.

content. The proposed scheme aims to provide the seamless playback for improving the QoE. To achieve this goal, we assign the different segment duration according to each quality level and request the extra segment from the server based on proposed rate adaptation scheme. Through the simulation results, the proposed scheme is proven to the seamless playback without buffer underflow. In the future work, we will analyze the proposed scheme in various network environments and implement the proposed system in real network.

REFERENCES

[1] M. Seufert, S. Egger, M. Slanina, T. Zinner, T. Hobfeld and P. Tran-Gia, "A Survey on Quality of Experience of HTTP

Adaptive Streaming," *IEEE Communications Survey and Tutorials*, vol. 17, no. 1, pp. 469–492, March 2015.

[2] ITU-R Recommendation BT.2020 "Parameter Values for UHD System for Production and International Programme Exchange," August 2012.

[3] D. M. Nguyen, L. B. Tran, H. T. Le, N. P. Ngc, and T. C. Thang, "An Evaluation of Segment Duration Effects in HTTP Adaptive Streamin over Mobile Networks," *in Proceeding of the 2nd National Foundation for Science and Technology Development Conference on Information and Computer Science (NICS '15)*, pp. 248-253, September 2015.

[4] C. Liu, I. Bouazizi, and M. Gabbouj, "Rate Adaptation for Adaptive HTTP Streaming," *in Proceeding of the ACM Multimedia Systems*, pp. 169-174, February 2011.

[5] S. Akhshabi, A. C. Begen, and C. Dovrolis, "An Experimental Evaluation of Rate-adaptation Algorithms in Adaptive Streaming over HTTP," *in Proceeding of ACM Conference on Multimedia Systems*, pp. 78-85, February 2011.

[6] T. C. Thang, Q. D. Ho, J. W. Kang, and A. T. Pham, "Adaptive Streaming of Audiovisual Content using MPEG DASH," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 1, pp. 78.85, February 2012.

[7] T. Y. Huang, N. Handigol, B. Heller, N. McKeown, and R. Johari, "Confused, Timid, and Unstable: Picking a Video Streaming Rate is Hard," *in Proceeding of ACM Coference on Internet Measurement*, pp. 225-238, November 2012.

[8] K. Miller, E. Quacchio, G. Gennari, and A. Wolisz, "Adaptation Algorithm for Adaptive Streaming over HTTP," *in Proceeding of the IEEE Packet Video Workshop*, pp. 173-178, May 2012.

[9] D. Suh, I. Jang, and S. Pack, "QoE-Enhanced Adaptation Algorithm over DASH for Multimedia Streaming," in Proceeding of IEEE Conference on Information Networking, pp. 497-501, February 2014.

Two-step Facial Images Deblurring With Kernel Refinement Via Smooth Priors

Jun-Hong Chen[1] and Long-Wen Chang[2]

Department of Computer Science

National Tsing Hua University

Hsinchu,Taiwan

Email[1]: jhchen.nthu@gmail.com

Email[2]: lchang@cs.nthu.edu.tw

Abstract— **Image deblurring is a challenging task in image processing. It is an ill-posed problem to estimate the unknown blur kernel and recover the original image from a blurred image. There are many methods for blurred natural images; however, few of them are able to perform well on blurred face images. Based on L_0 norm prior, we propose a two-step method for the images deblurring. The proposed method does not require any facial dataset to initialize the gradient of contours or any complex filtering strategies. In the first step, we combine L_0 norm prior with our local smooth prior to predict the blur kernel. With simple Gaussian filtering, we could maintain the smooth region in the latent image. In the second step, we refine the previous estimated kernel. In order to discard low intensity pixels that seemed to be noises on the kernel, we impose the sparsity on the kernel. Experimental results demonstrate that our proposed algorithm performs well on the facial images.**

Keywords-image debluring; kernel; smooth prior.

I.   INTRODUCTION

Blind deblurring is a tough task in image processing. It can be modelled as

$$B = x * k + n,\qquad(1)$$

where B is an observed blurred image, x is a latent image, $*$ is a convolution operator, k is a blur kernel, and n is a noise. Recovering a sharp image from the single blurred image B is an ill-posed problem. It can be solved by infinite set of pairs of the blur kernel and the latent image. Recent researches have a significant advance on image deblurring by iterative methods to estimate x and k. We can obtain a recovered image and an estimated kernel in each iteration. Usually, the temporarily recovered image is referred to as a latent image. The latent image is a clear and deblurred image. With edge information, the latent image and the blur kernel are estimated iteratively [3][5][8][10]. Shan et al. [2] adopted a sparse image prior via global and local priors. Krishnan et al. [6] used the L_1/L_2 prior on the high frequencies of an image. Xu et al. [9] used an unnatural L_0 prior to select high gradient edges iteratively. Kotera et al. [11] used straightforward maximum a posteriori method with $L_{0.3}$ heavy-tail prior. In this paper, we add a local prior to maintain smooth regions and refine the estimated kernel by the L_0 prior. In contrast to Pan et al. [14], which is focused on facial image deblurring, we predict the latent image without dataset supports and our experiments are better in image quality with appropriate

parameters. In Section 2, related works are addressed. In Section 3, our proposed method is introduced. Section 4 describes our experimental results and Section 5 is the conclusion.

II.   RELATED WORK

Image deblurring can be divided into two parts: blind deblurring and non-blind deblurring. The former is to recover a sharp image with only a given blurred image; the latter uses deconvolution to get a better result than the former. The challenge on facial images deblurring is that there is less texture in facial images. The existing methods use few edges in the blurred image to estimate a blur kernel. Cho and Lee [3] used an explicit filter as the shock filter to process the blurred image to select its real edges. Xu et al. [5] created a metric to measure the usefulness edges which is defined by local gradient information. Bae et al. [8] focused on the informative edges in patches. They combined gradient magnitude, the edges of straightness and usefulness edges as mentioned above. Other works have constraints on the sparsity of image gradients [2][6][9][11]. Shock filter has been used in [3] as a sharpened reference image for deblurring. However, due to severe blurring motion, recovering facial contour is difficult due to wrong edge selection. The gradient of heavy-tail distribution [1][2][4][7][11] property on face images might not work well because of less textures in those images. In other works, different sparsity priors [6], such as the normalized prior $\frac{\|x\|_1}{\|x\|_2}$ were introduced. According to Xu et al. [9], we can use it to recover the latent image with only several iterations. Without extra filtering such as shock filters, the optimization progress is faster than other methods. Recently, Pan et al. [14] proposed a deblurring method which uses similar face contours in their exemplar dataset for initial guess, and solve the objective function by the L_0 prior of the gradient magnitude. They collected hundreds of images and the gradients of those images to build their exemplar dataset. The success of the method is due to the facial global structure which has similar contours. However, an additional dataset is required to choose the initial guess. Different from the method of Pan et al. [14], we propose a two-step method to refine our results without dataset support. We focus on the smooth regions in face images and preserve the flat regions

by a local smooth term to alleviate the problem of ambiguous edge selection.

III. PROPOSED METHOD

The proposed method is shown in Figure 1. The blurred image is processed with the kernel estimation to obtain the initial kernel. Then, with the adjustment of the initial kernel based on its sparsity, we can get the final estimated kernel and the deblurred image. The kernel estimation is done with a coarse-to-fine approach. We estimate the latent image (x step) and the blur kernel (k step) iteratively in different scales. The coarse-to-fine method solves x by minimizing the objective function with a local smooth term from a low image scale to a high image scale. The estimated kernel at the last scale will be the initial kernel next round. In the kernel adjustment, the predicted kernel is denoised by using L_0 regularization to discard its low intensity pixels. The objective function with our local smoothness term is

$$\min_{x,z}\|x*k-B\|_2^2 + \lambda\|\nabla x\|_0 + \alpha\|\nabla x - \nabla z\|_2^2 \quad (2)$$

where x is a latent image, B is a blurred image, and k is a blur kernel, ∇ is a gradient operator of two directions (horizontal and vertical).

by a factor of 2. In this function, we only choose high gradient pixels in the blurred image which makes the latent image sparse. We set

$$z = \sim t \circ \hat{x}, \quad (5)$$

$\hat{x}$ is the latent image, which is filtered by a Gaussian filter with initial variance $\sigma = 0.3$ and increases over each scale, $\circ$ is a pointwise product (pixel to pixel), and $\sim$ is a NOT operator. t is a binary image indicating high gradients in the latent image (using w),

$$t = \begin{cases} 1, & w \neq 0 \\ 0, & w = 0 \end{cases},$$

We dilate t to avoid the influence on high gradient edge,

$$t = t \oplus m,$$

where $\oplus$ is a dilation operator, m is a square structuring element whose width is a half of estimated kernel size.

In our experiment, we set the initial kernel size for 3 pixel, and increases it over each scale by factor of $\sqrt{2}$. Thus, the z map we introduced could be thought of as the set of all pixels in the latent image with filtering, which is exclusive of the high gradient pixels in each iteration. According to Shan et al. [2], the smooth region in the latent image is still smooth after motion blur. The results with our smooth term are shown
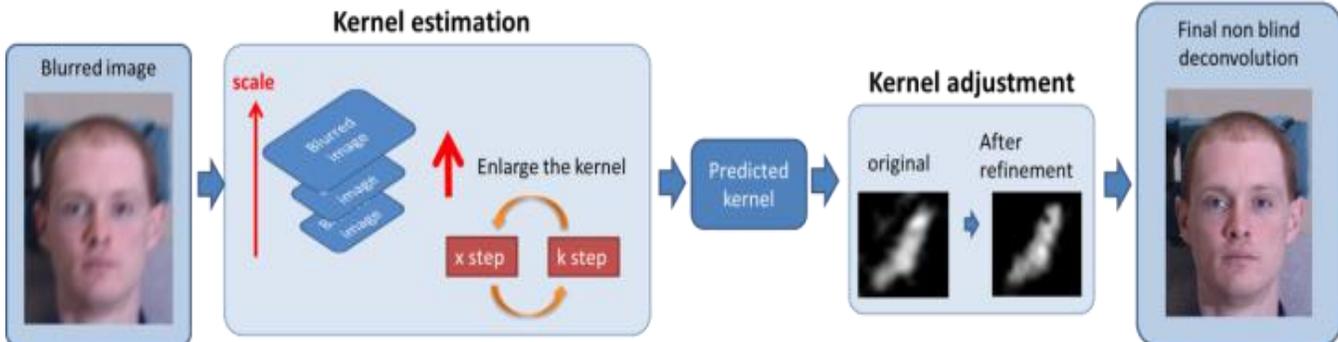

Figure 1. The flow chart of our proposed method

The first term represents the error between the observed blurred image and the estimated blurred image convoluted by the latent image and the kernel. The second term represents the constraint of the sparsity in the latent image. The third term is our proposed local smoothness term that we want the latent image to be close to the z map which indicate the smooth regions in the blurred image. As in Pan's [14], we can rewrite (2) as

$$\min_{x,w,z}\|x*k-B\|_2^2 + \beta\|\nabla x - w\|_2^2 + \lambda\|w\|_0 + \alpha\|\nabla x - \nabla z\|_2^2, \quad (3)$$

and

$$w = \begin{cases} \nabla x, & if\ |\nabla x|^2 \geq \frac{\lambda}{\beta}, \\ 0, & otherwise. \end{cases} \quad (4)$$

The variable w helps us to solve L_0 term. The larger β becomes, the closer the solution of (3) approaches (2). Because of the multiple-variable minimization of (3), we solve it alternatively by updating w, z and x independently. Given the latent image x, we obtain w by (4) with a threshold. Here, λ is a constant, and β becomes larger in each iteration
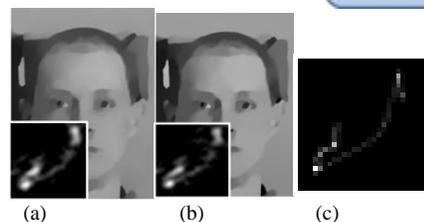

(a)   (b)   (c)

Figure 2. The latent image and estimated kernel recovered by our method with and without our priors of local smoothness. (a) Our local smooth prior ($\alpha = 0.0008,\ \lambda = 0.02$) (b) Without our local smooth prior ($\alpha = 0,\ \lambda = 0.02$) (c) original kernel

in Figure 2 (a). Minimizing (6) with respect to x, we obtain

$$\min_{x}\|x*k-B\|_2^2 + \beta\|\nabla x - w\|_2^2 + \alpha\|\nabla x - \nabla z\|_2^2. \quad (6)$$

According to Pan et al. [14], we can get

$$x = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(k)\mathcal{F}(B) + \beta\left(\overline{\mathcal{F}(\partial_x)}\mathcal{F}(w_x) + \overline{\mathcal{F}(\partial_y)}\mathcal{F}(w_y)\right) + \alpha\left(\overline{\mathcal{F}(\partial_x)}\mathcal{F}(z_x) + \overline{\mathcal{F}(\partial_y)}\mathcal{F}(z_y)\right)}{\overline{\mathcal{F}(k)}\mathcal{F}(k) + (\beta+\alpha)\left(\overline{\mathcal{F}(\partial_x)}\mathcal{F}(\partial_x) + \overline{\mathcal{F}(\partial_y)}\mathcal{F}(\partial_y)\right)}\right) \quad (7)$$

where $\mathcal{F}^{-1}(\cdot)$ and $\mathcal{F}(\cdot)$ are the Inverse Discrete Fourier Transform (IDFT) and Discrete Fourier Transform (DFT), and ∂_x and ∂_y are derivative operators in x and y directions, and $\overline{}$ is the complex conjugate operator. The objective function of the kernel estimation is

$$\min_{k}\|x * k - B\|_2^2 + \gamma\|k\|_2^2 \qquad (8)$$

where x is a latent image, B is a blurred image, and k is a blur kernel, γ is the parameter for regularization term on the kernel sparity. It can be solved by the conjugate gradient decent easily. Here, the data term is on gradient level according to Pan et al. [14], which has a stable kernel estimation can be written as

$$\min_{k}\|\nabla x * k - \nabla B\|_2^2 + \gamma\|k\|_2^2, \qquad (9)$$

where ∇ is gradient of horizontal and vertical direction. It is efficient to rewrite (9) in the matrix form. Thus, (9) can be rewritten as below:

$$(\nabla x \mathrm{K} - \nabla B)^{\mathrm{T}}(\nabla x \mathrm{K} - \nabla B) + \gamma \mathrm{K}^{\mathrm{T}}\mathrm{K}$$
$$= \mathrm{K}^T \nabla x^T \nabla x \mathrm{K} - \mathrm{K}^T \nabla x^T \nabla B - \nabla B^T \nabla x \mathrm{K} + \gamma \mathrm{K}^T \mathrm{K}, \quad (1)$$

where K is a convolution matrix referred to k. By the minimization of (10) with respect to K we can obtain

$$(\nabla x^T \nabla x + \gamma)\mathrm{K} = \nabla x^T \nabla B \qquad (11)$$

Then, (11) can be solved by the conjugate gradient decent. The coarse-to-fine kernel estimation is with (11) and estimation of the latent image is with (4), (5) and (7). The experimental observation of output kernel shows that there are still some low intensity pixels that seemed to be noise. The straightforward concept is using L_0 norm to let the kernel be sparse. Based on L_0 norm method, this problem involves the following terms

$$\min_{k}\|\nabla x * k - \nabla B\|_2^2 + \lambda\|k\|_0, \qquad (12)$$

Which represents the constraint on the sparsity of blur kernel. Since the function in equation (12) is a non-convex function, we add an auxiliary variable r into (12) and rewrite it as the previous case

$$\min_{k,r}\|\nabla x * k - \nabla B\|_2^2 + \lambda\|r\|_0 + \beta\|k - r\|_2^2. \ (13)$$

In each iteration, we alternatively solve subproblems with respect to each variable r and k,

$$r = \begin{cases} k, & if \ |k|^2 \geq \frac{\lambda}{\beta} \\ 0, & otherwise \end{cases} . \qquad (14)$$

Here, λ is a constant, and β increases by a factor of 2 in each iteration. The threshold of the intensity of kernel pixels is to discard the low value of pixels.

$$\min_{k}\|\nabla x * k - \nabla B\|_2^2 + \beta\|k - r\|_2^2. \qquad (15)$$

After the variable r is obtained, minimization of (14) can be rewritten as below:

$$(\nabla x^T \nabla x + \beta)\mathrm{K} = \nabla x^T \nabla B + \beta\mathrm{R}, \qquad (16)$$

where K is a convolution matrix referred to k, R is a convolution matrix referred to r.

We can solve (16) by conjugate gradient decent. The concept in the second step is based on the progressive sparsity of L_0 norm. Because L_0 norm is a non-convex function, we have to use additional variables and iterations to solve. Therefore, in the k step (kernel estimation process), we estimate the kernel with L_2 norm regularization which can be done easily by conjugate gradient decent. Compared to Pan et al. [14], we rewrite the L_0 method in coarse-to-fine approach. Without dataset support, we solve it in a single image. Compared to Xu et al. [9], the local smooth prior is

introduced, and we use two steps to refine our predicted kernel. Figure 2 shows the refinement of kernel estimation. We can see that Figure 3 (b) shows the kernel after refinement. It is much closer to the kernel of ground truth than Figure 2(a) the kernel before the refinement.
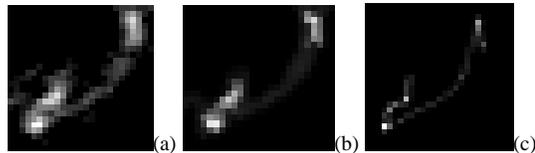

Figure 3. Refinement of kernel estimation (a) Before (b) After (c) Ground truth kernel

Finally, we obtain the refinement kernel and use the non-blind deconvolution provided by Pan et al. [14]: *deconvSps* function to get the recovering image.

IV.EXPERIMENTS

In our experiments, the parameters: $\alpha = [4e-4 \sim 64e-4]$, $\lambda=[0.02,0.03,0.04]$, and initial kernel size is 17×17. The dataset we test is provided by Pan et al. [14] and Levin et al. [11]. The ground truth images are the deconvolution of the blur images with the ground truth blur kernel. The non-blind deconvolution method is provided by Pan et al. [14]: deconvSps function, which is the same method we used in our estimated kernels to recover the sharp images for consistent measures. To validate our framework, we compare the visual quality of the recovered sharp latent image and kernel which is shown in Figure 4 without our local smooth prior and Figure 5 with and without our kernel refinement. Figure 6 shows the comparison of the results by various methods. The recovering method, deconvSps function, is the same provided by Pan et al. [14]. The image quality shows that the image recovered by our kernel looks softer and smoother. In addition, our local smooth prior and kernel adjustment could reduce the noise of the estimated kernel; that is, it is reasonable to fit two-step approach to improve the kernel. Also, 40 blur images provided by Pan et al. [14] are used. PSNR (Peak Signal to Noise Ratio) and SSIM (Structural Similarity) are the metrics between the images with ground truth kernel via non blind deconvolution and the deblurred images. Figures 7 and Figure 8 show that the comparison of PSNR and SSIM values for 40 images and our method has better performance.

IV. CONCLUSION

We present a new framework combined with L_0 norm prior for images which leverages those smooth regions and refines the kernel to get better results. The local smooth term is to maintain the smoothness in images. Without any dataset support, we use coarse-to-fine approach and perform well on face images.The better results could be credited to the flat region in these face images to alleviate the ambiguous edge selection.

REFERENCES

[1] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," ACM Transaction on Graphics, 25(3):787–794, 2006.

[2] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," ACM Transaction on Graphics, 27(3):73;1-73:10, 2008.

[3] S. Cho and S. Lee, "Fast motion deblurring," ACM Transaction on Graphics, 28(5), 145:1-145:8,2009.

[4] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding blind de-convolution algorithms," IEEE Transaction on Pattern Analysis AND Machine Intelligence, Vol. 33, No. 12, 2354-2367, 2011

[5] L. Xu and J. Jia: "Two-phase kernel estimation for robust motion deblurring," Proc. 11th European Conference on Computer Vision, 2010

[6] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure,". Computer Vision and Pattern Recognition, 2011.

[7] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Efficient marginal likelihood optimization in blind deconvolution, " Proc. IEEE Conf Computer Vision and Pattern Recognition, 2011.

[8] H. Bae, C. C. Fowlkes, and P. H. Chou, "Patch mosaic for fast motion deblurring," Proc. 11th Asian Conference on Computer Vision, 2012.

[9] L. Xu, S. Zheng, and J. Jia, "Unnatural l0 sparse representation for natural image deblurring," Proc. IEEE Conf Computer Vision and Pattern Recognition, 2013.

[10] L. Sun, S. Cho, J. Wang, and J. Hays, "Edge-based blur kernel estimation using patch priors," Proc. IEEE International Conference on Computational Photography, 2013

[11] J. Kotera, F. Šroubek, and P. Milanfar, "Blind deconvolution using alternating maximum a posteriori estimation with heavy-tailed priors," 15th International Conference on Computer Analysis of Images and Patterns, 2013.

[12] T. Michaeli and M. Irani, "Blind Deblurring Using Internal Patch Recurrence," Proc 13th European Conference European Conference on Computer Vision, 2014

[13] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring Text Images via L0-Regularized Intensity and Gradient Prior," Proc. IEEE Conf Computer Vision and Pattern Recognition,, 2014.

[14] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring face images with exemplars." Proc 13th European Conference European Conference on Computer Vision , 2014.
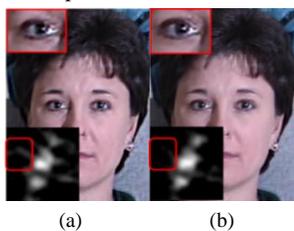
Figure 4 (a) without our local smooth prior. α=0, λ=0.02   (b) With our local smooth prior. α=16e-4, λ=0.02
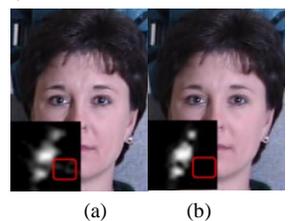

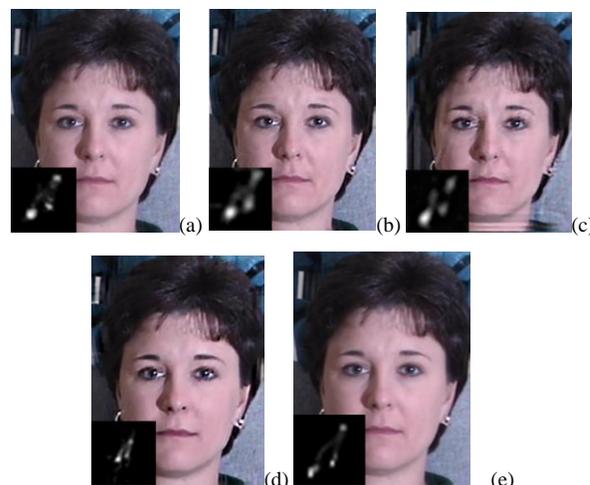Figure 5 (a) Without our kernel refinement (b) With our kernel refinement

Figure 6. Comparison of Image quality (a) Our method (b) Pan [14] (c) Kotera [11] (d) Xu [5] (e) Ground truth kernel with non-blind deconvolution
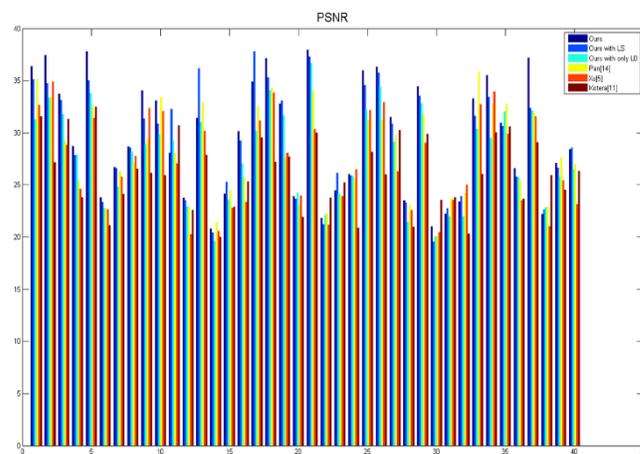

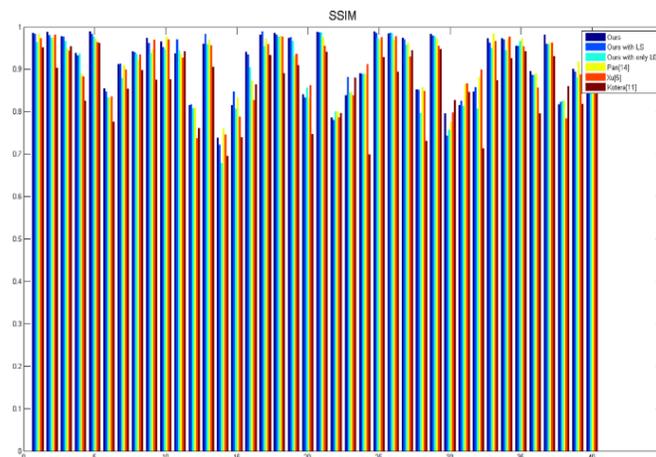Figure 7. Comparison of PSNR values for 40 blurred images

Figure 8. . Comparison of SSIM values for 40 blurred images