Optimal Allocation of 3G Budget for Smartphones Running Heterogeneous Applications

Saidur Rahman, Anika Anzum Prima and Md. Abdur Razzaque, *Senior Member, IEEE*
Green Networking Research (GNR) Group, Department of Computer Science and Engineering
Faculty of Engineering and Technology, University of Dhaka, Dhaka-1000, Bangladesh
Email: {sr.rifat, anikaprima13}@gmail.com, razzaque@du.ac.bd

***Abstract*—Significant growth in number of smartphone users and applications running on them has been observed in recent decade. The applications require diverse amount of data bandwidths based on their interactivities. Data hungry applications demand huge plan whereas a background application is satisfied with a minimum amount. Thus, smartphone applications should be allocated their required budget in such a way that resource wastage is minimized and user experience is maximized. In this paper, we develop a prioritized and dynamic budget allocation policy for ensuring optimal amount of budget allocation to each application and improve system performance. In this regard, we formulate a linear programming optimization function that maximizes the utilization and minimizes resource wastage. We also develop runtime monitoring technique for estimating future bandwidth utilization. Experimental results confirm that system performance goes up using proposed algorithm and proves effectiveness of the algorithm.**

I. Introduction

The recent years have observed exponential growth in the usage of smartphones and applications running on them. The diversity of the applications is increasing day by day with the rapid development of smartphone technologies as well as various wireless access technologies. smartphones are powerful enough to run heterogeneous applications concurrently. The unique combination of features makes smartphones extremely usable and useful for different purposes. smartphone applications provide diverse kinds of services besides simple voice communication. smartphones are featured with music player, high megapixel cameras, better navigators, diverse sensors and so on. It is being accepted that in the future smartphones will take over all the other digital devices in next years such as laptops, desktop computers and notebooks. It is revealed that many people use 3/4/5G capable smartphones which allow the users to access the Internet from almost anywhere at anytime. The next generation telecommunication standards provide cost efficient, high quality, wireless multimedia applications and enhanced wireless communications. It offers greater security features and high data transmission rate at a low cost. Today's smartphone's indispensable part is Internet centric applications. These applications have heterogeneous sensitivities to delay-deadlines to environmental changes and different bandwidth utilization. [1], [2], [3], [4]

smartphone applications with internet accessibility are in the heart of user's digital life. The applications upload or download data via Wi-Fi or 3G network. When a user can access Wi-Fi communication, all the data packets in the applications buffer are uploaded to the destination server through Wi-Fi communication regardless of its priority. But, if Wi-Fi signal is absent or too low to upload the important data packets, then the system can autonomously switch to 3G communication. Every application needs to be allocated limited 3G budget according to their bandwidth utilization. A significant amount of data from some low priority applications may easily blow through expensive 3G data plan and cause exhaustive use of constrained bandwidth resources. These applications can do a large amount of downloading or updating, e.g., weather updates, application updates, social networking application updates, etc.. These low priority and background applications can quickly chew up large volume of bandwidth from fixed data plan [5]. Consequently, all important or more sensitive applications will be deprived of uploading data or performing their tasks properly. Another problem is that a significant amount of data may remain unused at the end of the data plan for lack of proper distribution of budget among the applications. Therefore, it is essential to set a proper budget plan for smartphone applications that can apply smart policies to reduce the wastage of bandwidth resources as well as increase the user experiences. Resource provisioning is another key consideration. If an application is allocated less budget than it requires then that results in *under-provisioning*. On the other hand, if an application is allocated so high amount of budget that a significant amount of budget remains unused at the end of data plan period, that results in *over-provisioning*, as shown in Fig. 1. Our aim is to ensure efficient and systematic 3G budget utilization for each application for avoiding penalties incurred due to both over- and under-provisioning.

Efficient and effective utilization of bandwidth is a challenging task. Most of the state-of-the-art works [6], [7] did not consider dynamic budget allocation for smartphone ap-

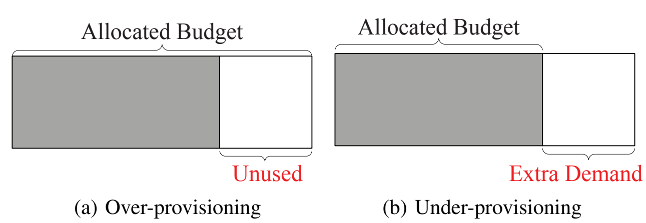


Fig. 1. Budget allocation penalties

plications. Bandwidth requirement is satisfied for constant bit rate and variable bit rate connection and connection blocking probability is kept low as well [7]. In [6], a heuristic solution to the problem of allocating budgets to sensitive and non-sensitive applications has been developed. However, it did not handle over- and under-provisioning problems while allocating budgets. Thus, they failed to make difference among applications with much diverse budget requirements, causing degraded performance in achieving better bandwidth utilization as well as many applications may be deprived of required bandwidth due to poor allocation policies.

In this work, we explore dynamic budget allocation policy which analyzes the budget usage behavior of each application and decides how to allocate resources for each application such as maximization of bandwidth will be ensured. Our proposed scheme of resource allocation ensures judicious amount of budget allocation for each application so that more important applications are not hampered performing their tasks. It also increases the overall bandwidth utilization for the applications.

The key contributions of this work can be summarized as follows:

- In this paper, we develop a prioritized and dynamic 3G budget allocation technique for smartphone applications.
- A bandwidth allocation function is formulated using linear programming optimization that maximizes budget utilization of all applications while minimizing allocation error as much as possible.
- A runtime monitoring and measurement scheme for estimating budget utilization has been developed using Weighted Average Usage Prediction (WAUP) method.
- We also recommend the amount of budget for future data plan using ARIMA (Autoregressive Integrated Moving Average) model. We carry out numerical evaluations to study the effectiveness of the proposed optimal budget allocation policy.
- The results show that the proposed model provides better performances than a number of state-of-the-art models.

The rest of the paper is organized as follows. Section II describes some of the works related to our topics of interest. The Section III presents system model. In section IV, we formulate optimization problem and propose dynamic budget allocation scheme. The Section V presents the result of performance evaluation and conclusions are drawn in Section VI.

II. Related Works

Allocating bandwidth among the competing users or devices is a challenging problem and it has been studied in the literature for many networks. In [7], the authors presented a utility based bandwidth allocation algorithm for multiple services in the heterogeneous wireless access networks consisting of WMAN, 3G cellular network and WLAN. Bandwidth is allocated to a new arrival connection in heterogeneous wireless environment depending on utility fairness. The researchers of [8] have proposed a smart bandwidth allocation algorithm based on smartphone users' personality traits and channel condition. Based on one user's data usage, the service provider could estimate this user's probability of each personality trait using diagnostic inference, and then based on predictive inference to calculate this user's usage of bandwidth in the future. The researchers of [9] have dealt with bandwidth disposition problem for heterogeneous networks. Their proposed method determines the amount of disposed bandwidth and upgraded or downgraded sequence of bandwidth is quantified by using Upgrade Rank or Downgrade Rank function.

The authors of [6] have introduced online 3G budget algorithm that decides which sensory data should be uploaded via 3G communication while others will be uploaded or downloaded later when Wi-Fi access point is encountered. Their optimization scheme ensures efficient 3G budget utilization but the algorithm causes large amount of computational overheads. Therefore, the approach is both computational resource and energy hungry. Also they have proposed a heuristic algorithm and the main focus of their proposed algorithm is to split overall 3G budget in each time cycle into two pieces: reserved budget and flexible budget. Sensitive applications use reserved budget and non-sensitive applications use flexible budget. If reserved budget runs out then sensitive applications take help from flexible budget. But this two-state classification (sensitive and non-sensitive) of the applications decreases the dynamicity and flexibility of bandwidth allocation. In addition to that the budget allocation strategies for heterogenous applications following their urgency have not been explicitly discussed and analyzed.

III. System Model

In this section, we present the system model for 3G budget utilization. We consider that a smartphone is connected with the Internet either using Wi-Fi access point or by 3/4/5G mobile Internet connection. The smartphone uses 3G bandwidth budget for urgent application usage whenever no Wi-Fi access point is available at nearby. The smartphone applications are allowed to buffer the data packets at local device till it is connected with any AP. In the case, the buffer space of the mobile phone is exhausted, it stops data collection process. When a user is in the range of a Wi-Fi access point, all the backlogged data packets in the buffer are uploaded to the destination server through Wi-Fi communication regardless of its priority. However, if Wi-Fi signal is absent or too low to upload the important data packets, then the system switches to 3G communication.

We assume that a user has fixed budget for 3/4/5G Internet connection (e.g., 3GB monthly, 1GB weekly package). The amount of the budget data plan for each of the applications is proportional to how much important the application is. That is, a real-time and interactive application needs more bandwidth and it may not tolerate significant delay; on the other hand, some low priority applications may be delayed and reduced amount of data budget can be allocated. In this work, we dynamically prioritize all the applications running in the mobile device by estimating bandwidth usage behavior of the applications. The more bandwidth an application uses, the higher it's priority is. We exploit autogregressive integrated moving average (ARIMA) formulae for estimating the runtime usage of resources by different applications and recommend a user the most appropriate amount of monthly data plan (to be discussed in detail in section IV-D).

We also assume that the budget allocation algorithm periodically runs every after t time. It tries to avoid overprovisioning as well as underprovisioning so as to maximize the resource utilization and application performance. Each application falls in one of the n application types with different priorities $p_1, p_2, p_3, \ldots, p_n$. In this case, we use higher values for higher priority applications.

IV. Proposed Model

In this section, we present the proposed dynamic budget allocation strategy for heterogeneous applications running in a smartphone. The proposed budget allocation policy dynamically expands or shrinks the amount of bandwidth allocated to different applications over time based on the usage behavior of the data plan. Our budget allocation optimizes the bandwidth resource utilization as well as reduces the penalties incurred due to over- and under-provisioning. We exploit Weighted Average Usage Prediction (WAUP) method to more accurately infer the bandwidth usage in future time intervals. We use Autogressive Integrated Moving Average (ARIMA) model for recommending appropriate amount of monthly data plan for a user.

A. Optimization Problem Formulation

The problem of optimal allocation of bandwidth to the mobile applications is translated as maximizing the utilization of resources while minimizing the penalties incurred due to over- and under-provisioning. And, this policy needs to be maintained for all applications in all allocation intervals. Therefore, the optimization function is a linear programming (LP) problem, defined as follows:

Maximize:

$$Z = \sum_{i=1}^{n} \sum_{t=1}^{T} (U_{i,t} - C_{i,t}) \tag{1}$$

subject to:

$$\begin{aligned} x_{i,t} &\leq r_{i,t}, & (2)\\ y_{i,t} &\leq f_{i,t}, & (3)\\ x_{i,t} + y_{i,t} &\leq r_{i,t} + f_{i,t} & (4) \end{aligned}$$

Here, $U_{i,t}$ is the resource utilization of application i at time interval t and $C_{i,t}$ is the corresponding over-provisioning penalty, if there is any. Given that the $r_{i,t}, x_{i,t}$ and $y_{i,t}$ are the amount of reserved budget, used reserved budget and used flexible budget for application i at time interval t, the utilization and penalties are defined as follows:

$$U_{i,t} = \begin{cases} 0 & \text{if } \frac{x_{i,t}+y_{i,t}}{r_{i,t}} > 1 \\ 1 & \text{if } \frac{x_{i,t}+y_{i,t}}{r_{i,t}} \leq 1 \end{cases} \tag{5}$$

$$C_{i,t} = \begin{cases} 0 & \text{if } \frac{x_{i,t}+y_{i,t}}{r_{i,t}} > 1 \\ \frac{(r_{i,t}-x_{i,t})}{r_{i,t}} & \text{if } \frac{x_{i,t}+y_{i,t}}{r_{i,t}} \leq 1 \end{cases} \tag{6}$$

We observe from Eq. (5) that the system performance decreases with the increasing usage of flexible budget $y_{i,t}$ and the Eq. (6) states that penalty increases with the gap between the amount of reserved budget $r_{i,t}$ and used reserved budget $x_{i,t}$. In summary, the more approriate amount of budget that we can allocate which just meets the requirement of an application, the more the system performance is increased and vice-versa. The constraints (2) and (3) are corresponding to bandwidth usage constraints for reserved and flexible budgets, i.e., usage must be bounded by the proportionately allocated amount for an application i. The constraint (4) states that the constraints (2) and (3) follow additive rule.

B. Budget Allocation Policy

The overall 3G budget B is split into two parts in each time cycle: *reserved* budget (B_1) and *flexible* budget (B_2). Initially ($t = 0$), their values are determined as follows:

$$B_1 = \alpha \times B, \tag{7}$$

$$B_2 = B - B_1, \tag{8}$$

where, α is a control parameter that determines how much of the total budget is to be kept in reserved portion. When the data plan period starts every application is allocated a certain amount of reserved budget based on their priority assuming all the applications have equal bandwidth usage for k time cycles. If n applications are running then,

$$r_{i,t} = \frac{p_i \times \bar{b}_{i,t}}{\sum_{i=1}^{n}(p_i \times \bar{b}_{i,t})} \times B_1, \tag{9}$$

where, p_i, $r_{i,t}$ and $\bar{b}_{i,t}$ are the priority, reserved budget and the estimated bandwidth usage of i'th application, respectively, within the time cycle. The detail estimation process of $\bar{b}_{i,t}$ is presented in Section IV-C. We assume that in the first time cycle t_1, an application i has used $x_{i,t}$ amount of data from the reserved budget. So, remaining reserved budget is $B_1 - \sum_{i=1}^{n} x_{i,t}$.

In the case, an application is run out of it's reserved budget within the current time cycle, then flexible budget is allocated to it from B_2. If $f_{i,t}$ denotes the extra budget requirement for i'th application in t time cycle then,

$$f_{i,t} = \frac{p_i \times T^{'}}{T \times \sum_{i=1}^{n} p_i} \times B_2 \tag{10}$$

If each application uploads or downloads $y_{i,t}$ amount of data using flexible budget then remaining flexible budget is $B_2 - \sum_{i=1}^{n} y_{i,t}$. So remaining total budget after t_1 time cycle,

$$B = (B_1 - \sum_{i=1}^{n} x_{i,t}) + (B_2 - \sum_{i=1}^{n} y_{i,t}) \tag{11}$$

This is the budget for next time cycle that means the assignment is additive. We now calculate total budget by adding remaining flexible and reserved budgets for the current time cycle. From the second time cycle, the reserved budget is calculated according to following equation:

$$B_1 = \frac{T^{'} + \alpha \times (T - T^{'})}{T} \times B \tag{12}$$

where, T is the budget validation time and $T^{'}$ is present time. The Eq. 12 helps us to dynamically update the reserved budget amount B_1 following the historical usages. It also minimizes the wastage of bandwidth later at the end of the data plan period. The control parameter α plays an important role to start with minimum reserved amount from the first day of data plan and to increase gradually. Therefore, it minimizes both the over- and under-provisioning penalties. The value of α depends on execution frequency of the budget allocation algorithm compared to the total data plan period. For performance evalution, we have set $\alpha = \frac{t}{T}$, where, t is the time interval of executing allocation algorithm.

C. Estimation of Budget Usage

We calculate budget usage ratio $b_{i,t}$ after each usage interval, t, as follows,

$$b_{i,t} = \frac{x_{i,t} + y_{i,t}}{r_{i,t} + f_{i,t}} \tag{13}$$

Thus, the Eq. 13 refers to how much of the allocated budget is used by an application i. We need to predict the allocated bandwidth budget usage of each application so as to infer the judicious amount of budget to be allocated in the upcoming time cycle. The possible amount of usage of the budget by an application in the next time cycle typically depends on its historical usage patterns. And the most recent usage behavior puts more impact on the future usage estimations. In this work, we exploit Weighted Average Usage Prediction (WAUP) method similar to WALI model [10], [11] that works as follows. The WAUP measures the average bandwidth usage of i'th application in the current time cycle as a weighted average of last m time cycles as follows:

$$\bar{b}_{i,t} = \frac{\sum_{j=1}^{m}(w_j \times b_{i,j})}{\sum_{j=1}^{m} w_j} \tag{14}$$

For weights w_j:

$$w_j = \begin{cases} 1 & \text{if } 1 \le j \le \frac{m}{2} \\ 1 - \frac{j - \frac{m}{2}}{\frac{m}{2}+1} & \text{if } \frac{m}{2} < j \le m \end{cases} \tag{15}$$

For $m = 8$, this gives weights of 1, 1, 1, 1, 0.8, 0.6, 0.4, 0.2 for w_1 through w_8, respectively where the most recent four samples are equally weighted.

D. Budget Recommendation using ARIMA

ARIMA(Autoregrassive Integrated Moving Average) is a common and effective method as one kind of time series prediction method. ARIMA($p,\ d,\ q$) models are first introduced by Box and Jenkins in 1970 [12] for purposes of modeling time series data. The model is the combination of autoregression and a moving average models. The full form of ARIMA can be written as [13], [14]

$$B_l^{'} = c + \phi_1 B_{l-1}^{'} + \cdots + \phi_p B_{l-p}^{'} + \theta_1 e_{l-1} + \cdots + \theta_q e_{l-q} + e_l \tag{16}$$

ARIMA($p,\ d,\ q$) can also be written as

$$B_l^{'} = c + \sum_{i=1}^{p} \phi_i B_{l-i}^{'} + \sum_{i=1}^{q} \theta_i e_{l-i} \tag{17}$$

or by using lag polynomial operator:

$$\bigtriangledown^d B_l^{'} \phi(L) = \theta(L) e_l \tag{18}$$

where,
$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - \cdots - \phi_p L^p$
$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3 - \cdots - \theta_p L^q$

$B_l^{'}$ is correlated normally distributed random variable, e_l is an uncorrelated Gaussian noise, θ_l is moving average coefficient and L is the lag operator.

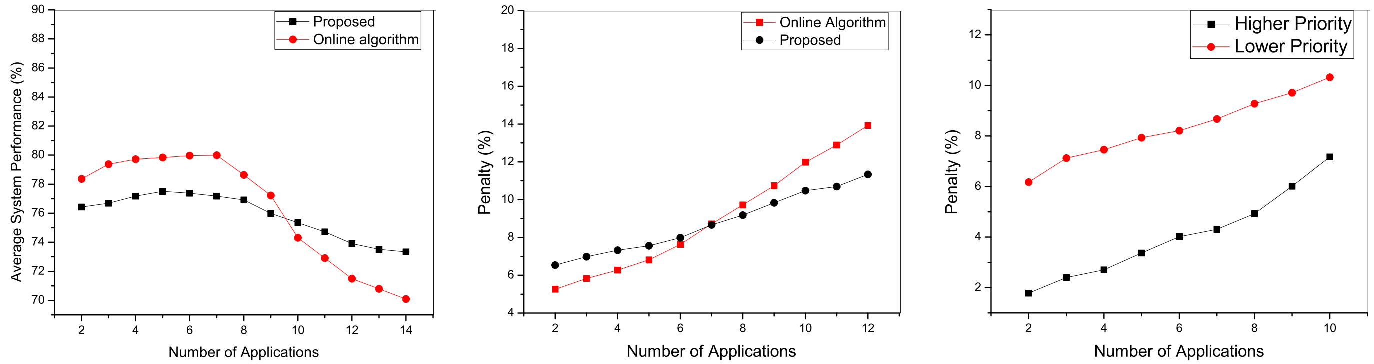
For predicting data plan period, ARIMA model is the better option for forcasting the data. Here, $B_l^{'}$ is the estimated data budget for lth data plan period. If the data is not stationary, then $B_l^{'} = B_l - B_{l-1}$.

V. Numerical Evaluation

In this section, we study the effectiveness of the proposed dynamic budget allocation policy compared to an online 3G budget algorithm [6] through numerical evaluations. We have implemented both the budget allocation algorithms using C++ programming language. We assume many applications are running in a mobile device and they have diverse bandwidth requirements. The number of applications that are active on a mobile device and it's monthly data plan are randomly chosen from a wide range of $2 \sim 12$ and $3 \sim 10$GB, respectively, with uniform distribution. The arrival and departures of applications are exponentially distributed. As a result, the duration for which an application keeps it active varies greatly from others. We also emulate that the mobile user does not use 3G data plan continuously for Internet accessibilities, rather sometimes it uses Wi-Fi access points for data transfer. The total data plan period is assumed to be 720 hours (i.e., 30 days) and the budget allocation algorithm execution time interval t is chosen 4-6 hours. For each of the graph data points, we run the program 20 times for different random inputs and take the average of the results.

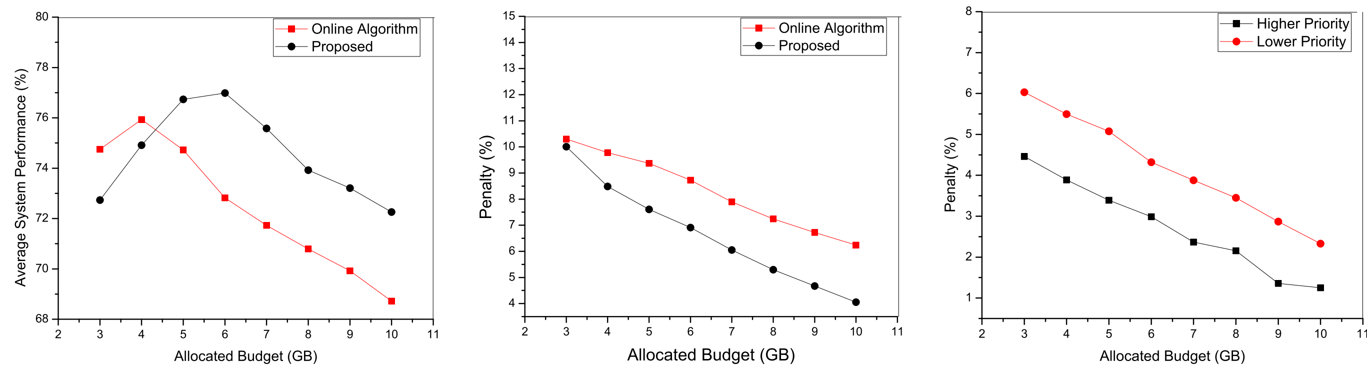
We have studied the following two metrics: *average system performance* and *penalty* for varying number of applications running on the smartphone. Eq. (1) defines system performance denoting the difference between resource utilization and over-provisioning penalty of each application in all time cycles. Our aim is to upgrade system performance by maximizing utilization and minimizing penalty. The *average system performance* is measured using Eq. (1) for all applications and then the average is taken for graph data points. The penalty is measured as the percentage of applications that could not be run due to shortage of bandwidth during the experiments. The average is taken for all time periods and all applications.

As shown in Fig. 2(a), the average system performance linearly increases with the number of applications in both the studied budget allocation algorithms. However, the performance of online 3G budget allocation algorithm starts



(a) System performance (b) Execution penalty (c) Execution penalty

Fig. 2. Performance studies for increasing number of mobile applications.



(a) System performance (b) Execution penalty (c) Execution penalty

Fig. 3. Performance studies for increasing amount of 3G bandwidth budgets.

decreasing when the number of applications is 8 and above. On the other hand, the proposed optimal budget allocation policy offers as high as almost 80% performance for higher number of applications. This happens because of its higher capability of accommodating diverse applications with different priorities and dynamically adjusting the bandwidth allocation to the applications following their historical usage pattern.

The graphs of Fig. 2(b) depict that the percentage of penalty increases with increasing number of applications but the penalty offered by proposed algorithm remains relatively low comparing to online 3G budget algorithm.

The graphs of Fig. 2(c) depict that the percentage of applications that are deprived of required amount of bandwidth allocation increases exponentially for lower priority applications. In this case, the results are caused by excessive underprovisioning penalty. However, for the higher priority applications, the percentage is significantly low, which is expected theoretically as well.

Fig. 3(a) shows that the average system performance offered by proposed algorithm grows with larger amount of budget but starts decreasing due to over-provisioning penalty when the allocated budget is 7 and above. However, the average performance offered by online algorithm remains low because of their incapability of proper distribution of budget among applications and handling over- and under-provisioning. Fig. 3(b) depicts that underprovisioning penalty decreases with larger amount of allocated budget as the applications are allocated judicious amount of budget. Fig. 3(c) depicts that underprovisioning penalty decreases with larger amount of allocated budget for both higher and lower priority applications.

VI. Conclusion

In this paper, we proposed dynamic bandwidth allocation algorithm for heterogeneous smartphone applications. The proposed data plan usage policy maximizes the utilization of all applications while minimizing over- and under-provisioning. We exploit the application's behavior and recommend future data plan after long term analysis. Experimental results confirmed that our proposed scheme gives optimal solution and potentially brings benefits to users. Our model also gives the better performance to distribute the budget among the applications where penalty will be minimized.

Acknowledgements

This work is supported by a grant for the "Innovative Project (2013-2014)" - "ICT Assisted Safe Driving for Mitigating Road Accidents in Bangladesh", funded by the Information and Communication Technology Division, Ministry of Posts, Telecommunications and Information Technology, Government of Bangladesh. Dr. Md. Abdur Razzaque is the corresponding author of this paper.

References

[1] Aaron Smith. U.s. smartphone use in 2015. In *The Smartphone Difference*. Pew Research Center, April 2015.

[2] Trent D. Buskirk and Charles Andrus. Smart surveys for smart phones: Exploring various approaches for conducting online mobile surveys via smartphones. In *Journal*. Survey Practice, 2012.

[3] Ahmad Rahmati and Lin Zhong. Studying smartphone usage: Lessons from a four-month field study. *IEEE Transactions on Mobile Computing*, 2012.

[4] Oliver Amft and Paul Lukowicz. From backpacks to smartphones: Past, present, and future of wearable computers. *IEEE Pervasive Computing*, 8(3):8–13, 2009.

[5] Statistics and market data on mobile internet & apps. http://www.statista.com/markets/424/topic/538/mobile-internet-apps/, (accessed July 16, 2015).

[6] Hengchang Liu, Shaohan Hu, Wei Zheng, Zhiheng Xie, Shiguang Wang, Pan Hui, and Tarek F. Abdelzaher. Efficient 3g budget utilization in mobile participatory sensing applications. In *INFOCOM*, pages 1411–1419. IEEE, 2013.

[7] Changqing Luo, Hong Ji, and Yi Li. Utility-based multi-service bandwidth allocation in the 4g heterogeneous wireless access networks. In *2009 IEEE Wireless Communications and Networking Conference, WCNC 2009, Proceedings, Budapest, Hungary, 5-8 April 2009*, pages 1915–1919, 2009.

[8] Junjie Chen, Qilian Liang, and Jie Wang. Bandwidth allocation based on personality traits on smartphone usage and channel condition. In *The Proceedings of the Third International Conference on Communications, Signal Processing, and Systems Part III*, pages 273–282, 2015.

[9] Hui-Min Huang and Ying-Hong Wang. Bandwidth management method for heterogeneous wireless network. *WTOC*, 7(4):267–276, April 2008.

[10] Sally Floyd, Mark Handley, Jitendra Padhye, and Jörg Widmer. Equation-based congestion control for unicast applications. In *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, SIGCOMM '00, pages 43–56, New York, NY, USA, 2000. ACM.

[11] Md. Abdur RAZZAQUE, Choong Seon HONG, and Sungwon LEE. Autonomous traffic engineering for boosting application fidelity in wireless sensor networks. In *IEICE Trans. Commun, Vol. E93-B*, pages 2990–3003. IEICE, November 2010.

[12] George Edward Pelham Box and Gwilym Jenkins. *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco, CA, 1970.

[13] Jie Wu Xin Jin, Yao Dong and Jujie Wang. An improved combined forecasting method for electric power load based on autoregressive integrated moving average model. In *2010 International Conference of Information Science and Management Engineering*, pages 476–480. IEEE, 2010.

[14] Mohamed Maaroufi Noreddine Citroen, Mohammed Ouassaid. Long term electricity demand forecasting using autoregressive integrated moving average model: Case study of morocco. In *1st International Conference on Electrical and Information Technologies (ICEIT)*, pages 59–64. IEEE, 2015.