Dynamic Demand Control with Differentiated QoS in User-in-the-Loop Controlled Cellular Networks

Rainer Schoenen[1,2], Halim Yanikomeroglu[1]

[1]Department of Systems and Computer Engineering, Carleton University, Canada
[2]Communication Networks (ComNets), Faculty 6, RWTH Aachen University, Germany

***Abstract*—Future cellular communications faces a number of challenges. One of the trends is the ever increasing demand for data rate due to smart mobile devices and laptop dongles with an estimated traffic growth of almost 100% per annum. Even with new cellular generation cycles every few years the same increase rate cannot be provided on the supply side. Neither anywhere nor anytime. The gap between supply and demand of wireless capacity will shorten and the conventional over-provisioning approach will not be possible anymore, especially during busy hours. The consequences are more frequent congestion situations with broken application traffic. The quality-of-experience will suffer as user expectations are high and steamed-up by advertising. An inadequate tariff system concentrating on flat-rates is also counterproductive for stability and energy-efficiency.**

In this paper the temporal user-in-the-loop (UIL) control approach is assumed. This user-centric model implements demand shaping by incentives in form of a dynamic usage-based tariff which adjusts based on the level of congestion in the busy hours. This is comparable to the smart grid operation principle. The novelty in this paper is the differentiated treatment for the exemplary service classes voice, video and data, for which new quantitative user response data is utilized. The control approach performance is calculated and results for stationary and dynamic scenarios are presented.

***Index Terms*—User-in-the-loop (UIL); demand shaping; congestion; tariff; QoS; sustainability; cross-layer.**

I. Introduction

4G and future cellular wireless networks promise to the end user ubiquitous access for all application rates and service classes with different QoS requirements at any time. In the same business flat-rates dominate the tariff structure and users expect to be able to have everything covered. This whole picture is between questionable and unrealistic. Flat tariffs have caps, many packages don't support all application ports or VoIP, users have to monitor their data usage, and traps and bill-shock management are business practice. Heavy users are feared by operators as they overuse the capacity. In the future, smart mobile devices and laptop dongles will be more and more common, leading to an enormous increase of requested data rate. Serious predictions estimate more than 100% increase of demand - a continuing trend since some years. Analysts therefore see a congestion problem within the next 10 years [1]–[4], no matter what inventions are made to increase the spectral efficiency by technical means. QoS for traffic classes with real time requirements, like voice and to a lesser extent streaming video, can only be guaranteed if the system is stable, i.e., in underload. This can only be achieved by over-provisioning bandwidth resources at all times of the day. Clearly, with all fluctuations during the day between night time and busy hours plus the exponential demand increase, over-provisioning is no longer a safe and viable strategy. The power consumption (green aspect) of always-on cellular base stations has also raised attention recently.

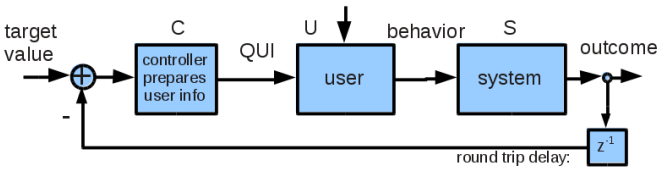

Fig. 1. User-in-the-loop (UIL): control of user and system [5].

In this paper, an aspect of the new user-in-the-loop (UIL) concept is treated. It is able to influence short-term user behavior and long-term patterns of use in order to use the user terminal (UT) at a better location or at a more suitable time. This can substantially improve the spectral efficiency and alleviate the traffic situation during the busy hours and in heavy congestion as previous work has shown [5]–[7]. Early previous work proposed usage-based pricing [8]–[11]. This paper investigates the temporal demand control of the user-generated traffic load by dynamic usage-based price rates with separated traffic classes for real-time (RT), non real-time (NRT) and best effort (BE) represented by voice/speech (S), video (V) and data (D). The stationary control properties are studied analytically and the dynamic properties by simulation. Three control approaches are compared and evaluated.

The paper organization is as follows. First the general model of UIL is introduced. Then the closed-loop control loop is defined and analyzed for multiple service classes. In the last section simulation results are presented.

II. User in the loop

The UIL concept was developed to control the user behavior in a wireless network in order to obtain a better spectral efficiency by convincing the users to move from one location to a better one [5] or to avoid traffic congestion by postponing session traffic out of the busy hours [7]. Depending on this impact dimension, the approach is called spatial or temporal UIL control. In both cases the user is within, and part of, a closed loop control system (Figure 1).

UIL extends the past assumption of the user being a traffic generating and consuming black box only. Instead the system

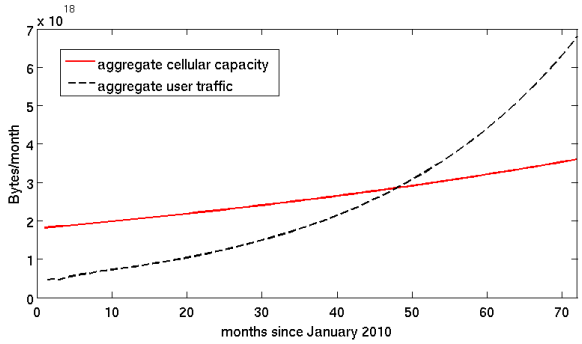

Fig. 2. Aggregate mobile traffic prediction $r^{(u)}(t)$ (u=unconstrained) and capacity limitation $\hat{R}(t) \approx r^{(t)}(t)$ (t=target rate). According to [4] demand will exceed the supply some day, if there are no measures taken to protect from this instability [7].

theoretic framework allows a control input to the user block, on which the user receives hints, incentives (and eventually penalties) in order to convince him to deviate from the default behavior (which is uncontrolled, open loop). A user within a closed control loop receives this control information (CI) in form of hints on the graphical user interface (GUI), e.g., a map and directions towards a better location, or the proposal of a better time to start his session (out of the busy hour). In Figure 1 this is called quantified user information (QUI).

The controller compares the system state (global and user-specific) with the target value (vector of goals), determines the severity and amount to control, translates it into user-perceivable CI and assumes that user and wireless system block react according to the control, so that the system output (performance) becomes as expected and determined, which in turn gives a low difference at the controller input.

Spatial UIL control is the first type of control. In this concept, a controller gives necessary location and incentive information to the user, and so it is expected that the user voluntarily changes his current location from point A to B, where a better signal quality is known from a database of statistically processed past channel measurements. The user moves from A to B with probability p_M which depends on the distance and the given incentive utility. Recent survey results are available to quantify this number based on given parameters [12]. The advantage is a substantially improved spectral efficiency $\bar{\gamma}$ and thus energy-efficiency. The spatial UIL control has been treated in [5], [6], [13]. Spatial control becomes even more relevant when it comes to femtocells and indoor navigation, where it is highly recommended to guide the user closer to the next hotspot. In this paper the focus is on temporal UIL control.

A. Temporal Control

The demand increase in cellular networks is fueled by the flat rate pricing policy which is dominating the market these days. Flat rates are favored by users only because of convenience and risk management, in order to avoid the risk of surprisingly exploding bills. Operators' business units like the customer binding of such contracts: money is flowing in from customers even if the data service is not used very much in some cases, which holds for the majority of users. Unfortunately, the top 20% of the users account for 80% of the traffic. Flat rates (especially the unbounded flat rates) promote heavy-tailed traffic distributions and in the long term the spiral of growth leads to a fast exponential demand increase [3], [4]. Figure 2 shows the total estimated traffic over the next years according to these forecasts and obviously there will be a congestion problem at some point in the future because wireless capacity cannot grow adequately with 100% per year. Flat rate pricing with a cap is only a transitional solution as it can only (open-loop) control traffic on a monthly basis.

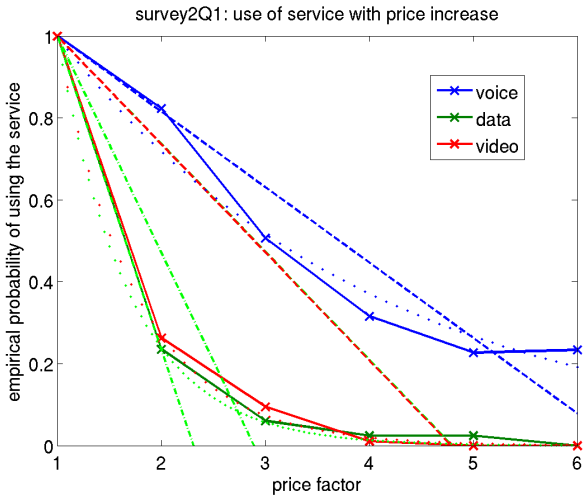

Fig. 3. Reaction to price increase $p(\chi)$ differentiated by service [12]. The linear and exponential fits are added as dashed and dotted lines, respectively, according to the fits in Eq. 1. Voice is less elastic than data or video.

In temporal UIL a fully dynamic usage-based pricing is suggested [7]. In the default mode the dynamic price is displayed on a user terminal (UT) before each (financially significant) application session transaction so that user can decide to use or not to use the service at the current time, location and price. In an advanced mode, an agent or manager software on the UT will act on behalf of the user, knowing his preferences (either by static settings in options or machine learning). The main idea is clear - the user will generate less traffic when the session price goes up. Recent survey results [12] indicate numbers how this would work. As a result the pricing method will change the user behavior and the traffic similar as it will do in future electricity tariffs and smart-grid applications. In the early UIL work [7] the user behavior is assumed as a linear with constant elasticity, but since the survey [12], more elaborate functions can be assumed, separated by the traffic classes for voice (S), video (V) and data (D). For the analysis the following exponential user responses are assumed (Figure 3) (alternatively linear):

$$\begin{aligned}
p_{voice}(\chi) &= e^{-\eta_S \cdot \chi} = e^{-0.330 \cdot \chi}; \quad (i.e., \eta_S = 0.330) \\
p_{data}(\chi) &= e^{-\eta_D \cdot \chi} = e^{-1.429 \cdot \chi}; \quad (i.e., \eta_D = 1.429) \\
p_{video}(\chi) &= e^{-\eta_V \cdot \chi} = e^{-1.304 \cdot \chi}; \quad (i.e., \eta_V = 1.304) \\
p^{(lin)}_{voice}(\chi) &= 1 - 0.18 \cdot \chi \\
p^{(lin)}_{video}(\chi) &= 1 - 0.73 \cdot \chi \\
p^{(lin)}_{data}(\chi) &= 1 - 0.52 \cdot \chi. \qquad (1)
\end{aligned}$$

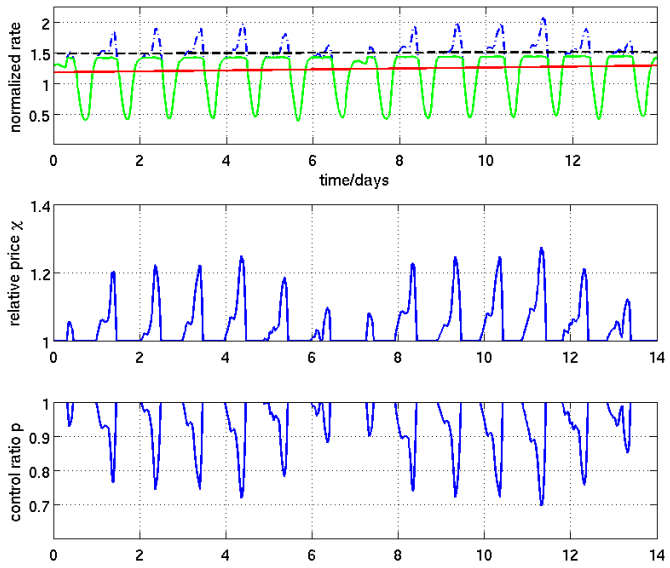

Fig. 4. UIL temporal control in times of predicted congestion during the busy hours. For this result only one traffic class was assumed (video). The displayed 14 days represent typical traffic week days from a Sunday to a Saturday. The blue dash-dot line is the unconstrained traffic demand, with average shown by the red line. The black dashed line is the capacity of the system. The green line is the rate after using UIL temporal control. The controller calculates the normalized price increase χ. The users answer with a demand reduction given by the control ratio p.

III. The Temporal UIL Controller

The controller **C** is a part of the management functionality in the network. It prepares the data to display to the user in a user-friendly and informative way. In Figure 1, **C** compares a target value (maximum utilization $u^{(t)}$ or rate $r^{(t)}$) with network measurements of the controlled rate $r^{(c)}$ (demand shaping). Note that this loop contains all users in a cell and all users get to see the same price rate π_C per class of service C (the unit of π_C is \$/bit but displayed in a user-understandable way). All variables here denote aggregates over all users and services. An example for the nominal prices $\pi_C^{(N)}$ is 8 ct/MB for voice, 4 ct/MB for video, 2 ct/MB for data, used later in Section IV.

The control ratio is defined as $p = r^{(c)}/r^{(u)}$ where $r^{(c)}$ is the controlled output rate and $r^{(u)}$ is the uncontrolled output rate (assuming a regular price level). An alternative interpretation for p is the proportion of users that do not change their original demand, while $1 - p$ of the users react and do not trigger the data transmission. The controller knows the error $\epsilon(\tau) = r^{(t)}(\tau) - r^{(c)}(\tau)$, for each time step τ. The control only needs to act in congestion, i.e., when $r^{(u)} > r^{(t)}$, else π_C stays on the default level (saturation or limiter block).

To reduce the uncontrolled traffic load $r^{(u)}$ to $r^{(t)}$, the control ratio p must be chosen as $p = min(r^{(t)}/r^{(u)}, 1)$. Then, depending on the pricing model [7], an adaptation of the pricing parameter is needed. Here we assume a pricing Π proportional to the usage volume v using π as the dynamic price=tariff rate. So Eq. 2

$$\Pi = \pi \cdot v = \cdot \int_{t=0}^{T_{month}} \sum_{\phi=1}^{\Phi} \pi_\phi \cdot r_\phi^{(c)}(t) dt \qquad (2)$$

determines the sum price for all user initiated session flows $\phi \in [1, \Phi]$ with class $C(\phi)$ and the current rate $r_\phi^{(c)}(t)$ and

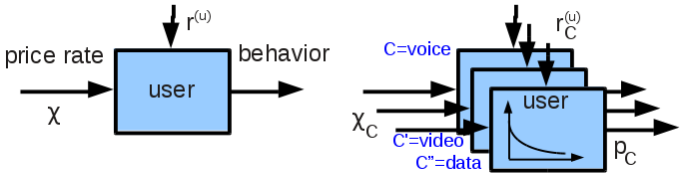

Fig. 5. The user box takes the unconstrained demand $r_C^{(u)}$ per traffic class C and is controlled by χ_C and generates/consumes real traffic with a rate of $r_C^{(c)} = p_C \cdot r_C^{(u)}$. There is also a system response in time due to the constant price assumption for the length of a flow per session.

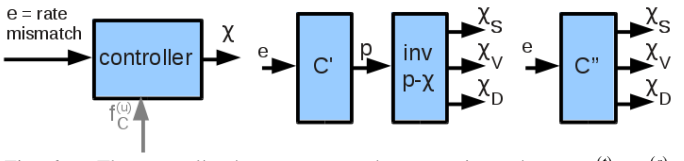

Fig. 6. The controller box processes the rate mismatch $e = r^{(t)} - r^{(c)}$, optionally with knowledge of the traffic class proportions $f_C^{(u)}$.

dynamic price π_ϕ which stays constant during the duration of a flow ϕ and is fixed at flow setup time $T_{0\phi}$ with the current tariff of the service class $C = C(\phi)$, $\pi_{C(\phi)}(t = T_{0\phi})$. We use χ_C as normalized price increase, per class C, with a default of 0.

$$\pi_C = \pi_C^{(N)} \cdot (1 + \chi_C). \qquad (3)$$

The price information π is known to the user block, so that the control loop is closed. The user reacts stochastically but in total over all users the reaction $p_C = f(\chi_C)$ leads to the decreased demand $r_C^{(c)} = p_C \cdot r_C^{(u)}$ per class C. See the example in Fig. 4. In a real world scenario the user behavior is assessed and improved (Kalman filter) in the live cellular system by gathering statistics on his conditional accept/deny pattern.

A. Dynamic Properties

For the purpose of demand shaping, there are two time scales. One is the long-term control to avoid traffic to exceed the capacity (Figure 2), and the other is a short-term control within hours, minutes or seconds to avoid short-term congestion situations due to busy hour peaks and bursty applications.

The user box is modeled as shown in Figure 5. The stationary (DC) input/output response the the input χ is a control ratio p which is estimated by Eq. 1 for simulation purposes in this paper but in the field would be estimated with the support of feedback from the operating system of the UT. The unconstrained traffic demand $r_C^{(u)}$ is divided into traffic classes C with $r_C^{(u)} = f_C^{(u)} \cdot r^{(u)}$ and projected proportions for 2020 [4] of

$$f_S^{(u)} = 10\%, f_V^{(u)} = 59\%, f_D^{(u)} = 31\%. \qquad (4)$$

The dynamic impulse response of the user block is determined by the statistics of the length of a session. These statistics are easy to obtain. A typical number for a website session duration is $\tau_D = 300\ s$ on a desktop and $\tau_D = 60\ s$ on a smartphone [14], $\tau_S = 100\ s$ for a mobile voice call and $\tau_V = 228\ s$ (3.8 minutes) for an online video. For the purpose of this study assuming an exponential distribution is sufficient. So the impulse response is modeled as

$$u_C(t) = p_c \cdot e^{t/\tau_C} \multimap \mathscr{L}\{u_C(t)\} = U_C(s) = p_C/(s \cdot \tau_C + 1). \qquad (5)$$

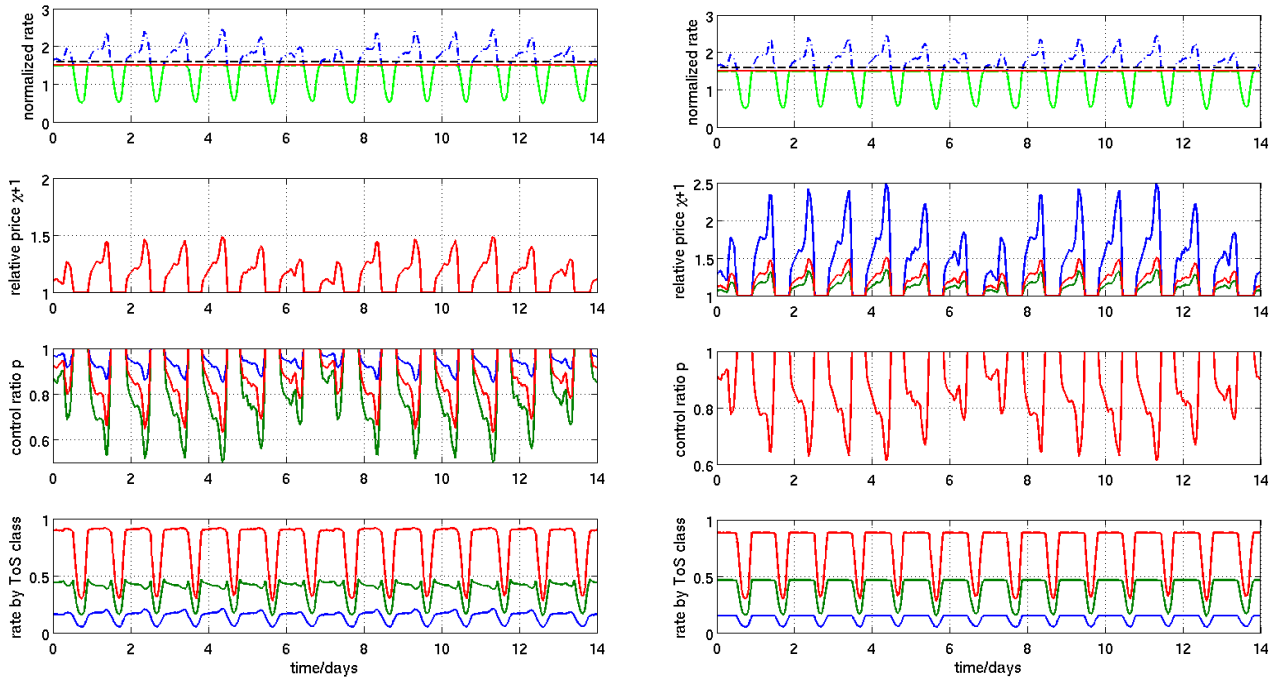

Fig. 7. UIL temporal control in total (top graph) and for all three service classes (blue=voice, red=video, green=data) for the control method M1="same relative price" left and M2="same control ratio" right. In both cases the sum rate is controlled below the target capacity. Only the proportions between voice, video, data differ. The dynamic transient behavior works well. All rate measurements are low-pass filtered to eliminate packet arrival fluctuations.

The controller must process the measured rate mismatch (error e) and generate the dynamic price rate indicator χ_C for each class C (Figure 6).

The control works well with a proportional+integral (PI) controller, the dynamics of which are given in continuous time:

$$\chi(t) = K \cdot (e(t) + \frac{1}{T_i} \int_0^t e(\tau) d\tau) \multimap\!\bullet\, \mathscr{L}\{c\} = C(s) = K \cdot \frac{s + 1/T_i}{s}. \tag{6}$$

Due to the relatively slow session load changes, $K = -\frac{1}{2}$ and $T_i^{-1} = 20/min$ work well in practice, implemented in a discrete-time version with a sampling time of $T_s = 60s$ and all traffic rates normalized to a capacity of 1 (corresponds to a load of $\rho = 1$). Measurements must be low-pass filtered to eliminate short-term rate fluctuations caused by the packet arrival stochastic process.

The system block S is responsible for multiplying p with $r^{(u)}$ to get the traffic rate $r^{(c)}$ and is therefore not a time-invariant function. The system is not a linear and time invariant (LTI) system. However, within the operating regions investigated here ($0 \le \rho \le 5$) this was manageable. The total system transfer function from in the Laplace-domain is $A(s)$ and the transfer function from $r^{(u)}$ to $r^{(c)}$ is $B(s)$:

$$A(s) = \frac{C \cdot U \cdot S}{1 + C \cdot U \cdot S}; B(s) = \frac{S'}{1 + C \cdot U \cdot S}. \tag{7}$$

Dynamic simulation results are provided in Figure 7.

IV. Results for class aware demand shaping

Figure 6 shows three different ways of controlling the three service classes. Method M1="same relative price" (left), M2="same control ratio" (middle) and M3="prioritized" (right). M1 simply controls as described above and doesn't care about service classes, i.e., the same normalized price rate χ is calculated for all classes. As the elasticity is different, this has a different effect on each classes traffic (Figure 7 left).

M2 aims at reducing each traffic by the same relative amount p. This control ratio is calculated by the controller ($K > 0$) first and then translated to different χ_C by taking the inverse of the user response function (Eq. 1) by

$$\begin{aligned} \chi_{voice}(p) &= -ln(p)/0.330 \\ \chi_{data}(p) &= -ln(p)/1.429 \\ \chi_{video}(p) &= -ln(p)/1.304. \end{aligned} \tag{8}$$

This can be adjusted by an adaptive function learning from statistics of real user responses, but it appears that exact numbers are not critical for the function and stability of the control. See the dynamics in Figure 7 right.

M3 is a more advanced control which calculates χ_C directly. The purpose is to reduce the cheapest traffic first until its price rate π_D (not normalized) reaches the next level of π_V, at which point both are controlled synchronously (with a relative gap of β kept between them). When π_V reaches π_S at some point, then all of them are controlled synchronously. This can best be seen in the second graph of Figure 10.

The three Figures, 8, 9, and 10 display the performance of the methods M1, M2 and M3. Note that the operating region of $\rho = \lambda/\mu = r^{(u)}/\hat{R}$ on the x-axis is beyond all numbers known from traffic and queueing theory, which requires $\rho < 1$ for stability. Here this $\rho \ge 1$ is the deep demand overload situation. The UIL closed loop control is used to control the traffic back to a level around $r^{(t)}(t)$ which corresponds to $\rho^{(t)} = 0.95$. This can also be interpreted as a "soft" connection admission control (CAC), because connections and sessions are accepted if the user is willing to pay the higher price for it.

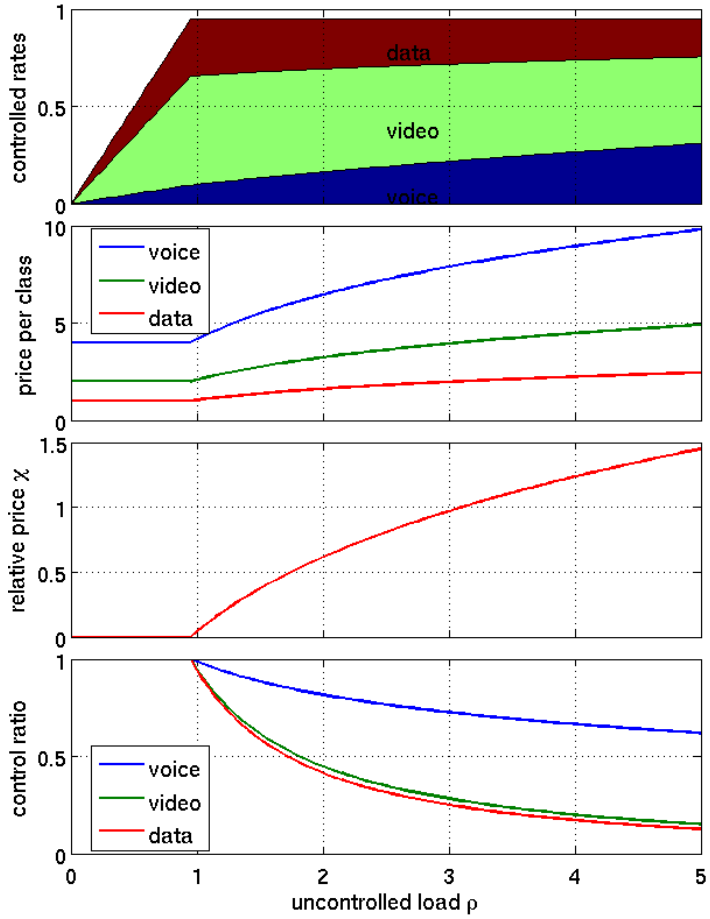

Fig. 8. Traffic control for "same price rate" rule (M1).

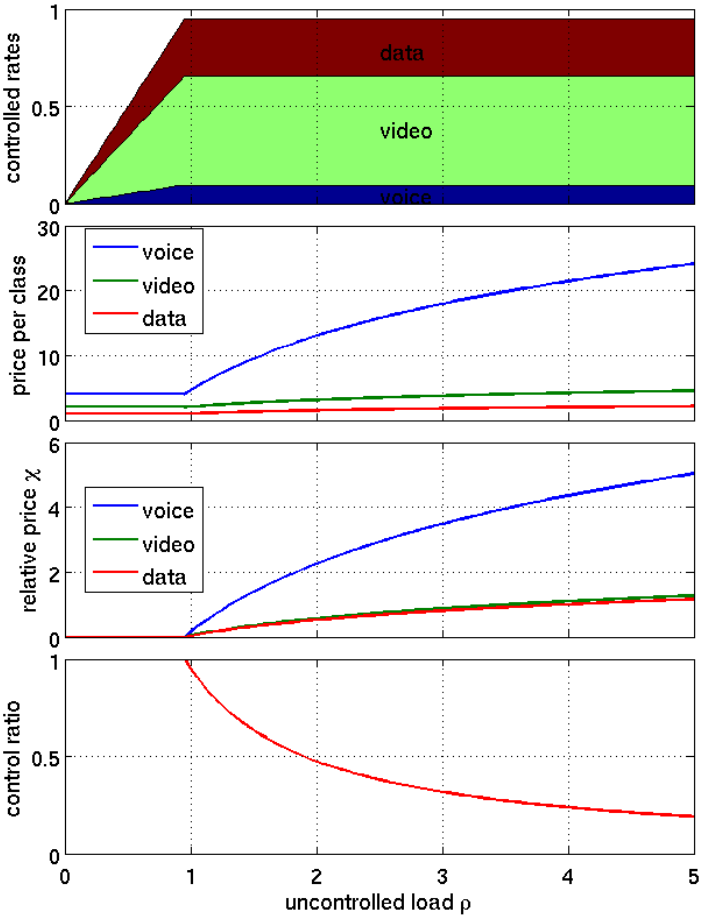

Fig. 9. Traffic control for "same control ratio" rule (M2).

M1 in Figure 8 is the simplest case, where all class dependent prices are multiplied with the same factor (graph 3). Therefore the user response is different per class (graph 4). In effect (graph 1), voice traffic is accepted more and more, even in deep demand overload. Note that the real voice traffic load is $\rho_S \approx 0.3$ even at $\rho^{(u)} = 5$. which is no problem to handle with all QoS requirements in a queueing system. The proportion of video and data is naturally reduced.

M2 in Figure 9 applies the same control ratio to all classes (graph 4). Because of the different user response, the price indicator χ_C must be calculated per class C (graph 3). All traffic classes are reduced by the same amount (graph 4), so in effect (graph 1), all classes maintain their proportion independent of the overload intensity.

M3 in Figure 10 applies the priority rule. Control ratio p_c and χ_C are calculated such that the cheapest service is modified first, as much as possible, and the order of services (by absolute price rate) is still preserved (graph 2). In effect (graph 1), voice traffic is only touched minimally in extreme demand overload and only because users are willing to pay a certain price also for video and data. This assumption is based on the surveyed data and may differ in reality, but the principle works with the algorithm explained here.

A. Stationary Analysis and Controller Calculations

In stationarity Eq. 7 can be simplified with $s \to 0$ and $e = 0$ ($r^{(t)} = r^{(c)}$) can be assumed due to the I-component of the controller. For M1, the stationary solution can be found (with $\Theta := r^{(u)}/r^{(t)}$) by

$$r^{(t)} = r^{(c)} = r^{(u)} \cdot \sum_{C=1}^{3} f_C^{(u)} \cdot p_C(\chi_C) \tag{9}$$

$$0 = \Theta \cdot \sum_{C=1}^{3} f_C^{(u)} \cdot e^{-\eta_C \cdot \chi} - 1 \tag{10}$$

The solution χ can be obtained by finding the zeros of Eq. 10. There is exactly one zero when $\Theta \geq 0$ and the function decreases monotonically from $+\Theta$ to -1. For a linear user response model (Eq. 1 below) χ can be explicitly calculated:

$$0 = \Theta \cdot \sum_{C=1}^{3} f_C^{(u)} \cdot (1 - \eta'_C \cdot \chi) - 1 \tag{11}$$

$$\chi = (1 - 1/\Theta) / \sum_{C=1}^{3} f_C^{(u)} \cdot \eta'_C \tag{12}$$

For model M3 also Eq. 9 has to be solved. The χ_C are different though. A numeric solution is possible by a distinction of cases whether the constraint in Eq. 13 affects a) no class, b) only D, c) D+V, d) all classes.

$$\pi_S \cdot (1 + \chi_S) \geq \beta \pi_V \cdot (1 + \chi_V) \geq \beta^2 \pi_D \cdot (1 + \chi_D) \tag{13}$$

Substituting $\chi_V = (1 + \chi_D) \cdot \beta \pi_D / \pi_V - 1$ and $\chi_S = (1 + \chi_D) \cdot \beta^2 \pi_D / \pi_S - 1$ into Eq. 9 gives lengthy terms in form of $f(\chi_D) = 0$ which can be solved by finding the zero. All figures 8,9,10 have been obtained by this analysis.

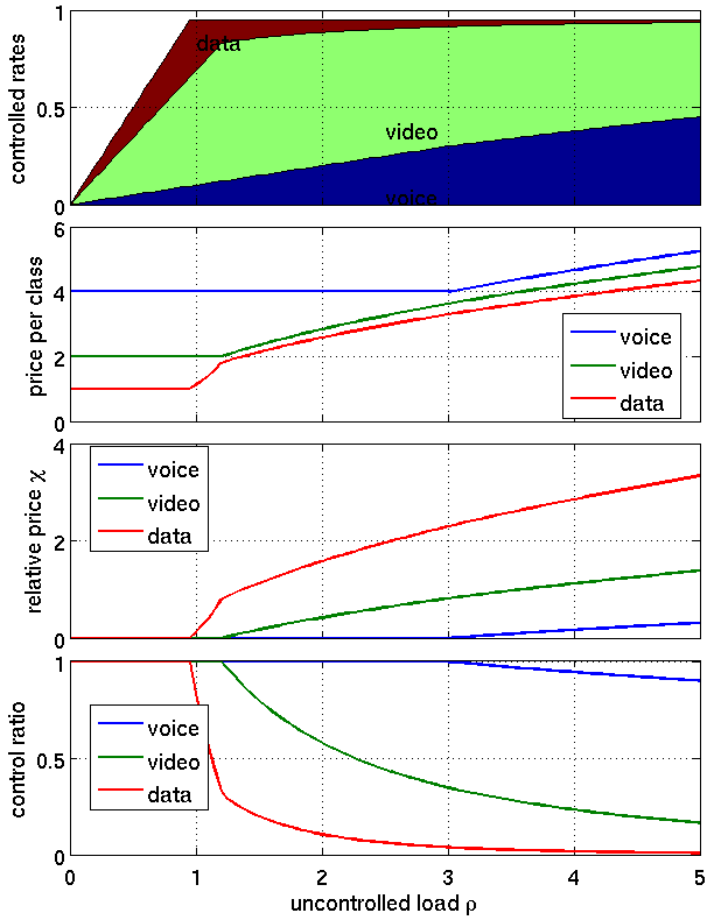

Fig. 10. Traffic control for "priority price" rule (M3, $\beta = 1.1$).

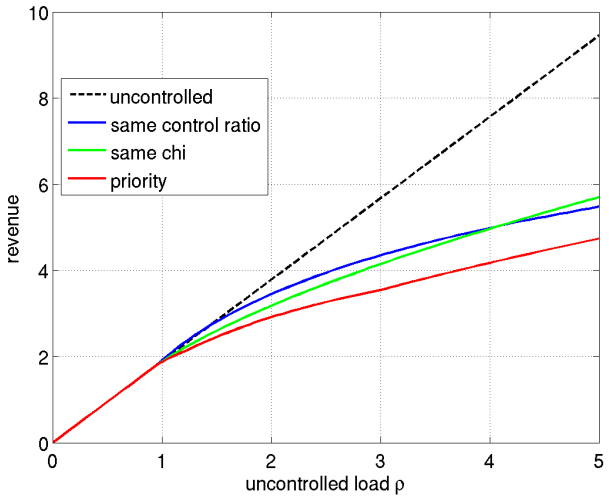

Fig. 11. Revenue comparison: Multiplying price rates with controlled data rates, this graph shows the total revenue rate [$/*time*] depending on the demand overload. The three methods M1,M2,M3 perform differently and the priority scheme M3 imposes the least financial burden on the customer, but also the lowest revenue for the operator. Clearly the revenue can be increased by operating in virtual overload, and between $1 \leq \rho \leq 2$ the revenue is close or even above the regular revenue, i.e., if the system were not in congestion and all traffic was carried (uncontrolled).

V. Conclusion

This paper treated the user-in-the-loop (UIL) control for differentiated services with independent tariff plans and user response. Given recent quantitative user response data, a class-aware control loop was constructed and three different controller methods were proposed and compared to handle demand overload situations. Such situations are more and more likely in future cellular scenarios given the heterogeneous nature of traffic in time and space as well as the limited wireless resources which are location-dependent and are already used to the limit of the spectral efficiency. The proposed temporal UIL control resolves congestion situations in the busy hour and manages to let users reduce their rate of video traffic and data traffic (downloads) because of the increased dynamic price. This is comparable to the intention of the smart grid, with the additional feature of service classes. Results show that it is possible to control traffic into a stable operating point and this is robust against inaccuracies of user properties and nonlinear effects. The three control methods show different treatment of the classes between voice and video+data. An economic outlook given in Figure 11 can be used to find the point at which operators should invest into new infrastructure.

The green aspect is that avoiding certain traffic at certain times can postpone infrastructure investments, therefore saving energy and operators' money. Having a price for each transaction (like watching a video for $1) leads users to be green-aware, which else is awkward with flat-rate tariffs.

References

[1] "UMTS Forum report 44 - mobile traffic forecasts 2010-2020," http://www.umts-forum.org/, UMTS Forum, Tech. Rep., Jan 2011.

[2] "Cisco visual networking index: Global mobile data traffic forecast update, 2010–2015," Cisco Systems Inc., White Paper, February 1, 2011. [Online]. Available: http://www.cisco.com/en/US/solutions/

[3] "2010 mobile internet phenomena report," Sandvine Inc., White Paper. [Online]. Available: http://www.sandvine.com/downloads/documents/

[4] "Mobile broadband capacity constraints and the need for optimization," Rysavy Inc., White Paper, February 2010. [Online]. Available: http://www.rysavy.com/Articles/

[5] R. Schoenen, H. Yanikomeroglu, and B. Walke, "User-in-the-loop: Mobility aware users substantially boost spectral efficiency of cellular OFDMA systems," *IEEE Communications Letters*, vol. 15, no. 5, pp. 488–490, May 2011.

[6] R. Schoenen, "On increasing the spectral efficiency more than 100% by user-in-the-control-loop," in *Proceedings of the 16th Asia-Pacific Conference on Communications (APCC)*, Auckland, October 2010.

[7] R. Schoenen, G. Bulu, A. Mirtaheri, and H. Yanikomeroglu, "Green communications by demand shaping and User-in-the-Loop tariff-based control," in *Proceedings of the 2011 IEEE Online Green Communications Conference (IEEE GreenCom'11)*, Online, 2011.

[8] C. Saraydar, N. Mandayam, and D. Goodman, "Efficient power control via pricing in wireless data networks," *Communications, IEEE Transactions on*, vol. 50, no. 2, pp. 291 –303, Feb. 2002.

[9] J. Altmann and K. Chu, "How to charge for network services - flat-rate or usage-based?" *Computer Networks*, vol. 36, no. 5-6, p. 519, 2001.

[10] C. Courcoubetis, F. Kelly, V. Siris, and R. Weber, "A study of simple usagebased charging schemes for broadband networks," *Telecommunication Systems*, vol. 15, pp. 323–343, 2000.

[11] J. MacKie-Mason and H. Varian, "Pricing congestible network resources," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1141–1149, Sep. 1995.

[12] R. Schoenen, G. Bulu, A. Mirtaheri, T. Beitelmal, and H. Yanikomeroglu, "Quantified user behavior in user-in-the-loop spatially and demand controlled cellular systems," in *Proceedings of the European Wireless*, Poznan, 2012.

[13] T. Beitelmal, R. Schoenen, and H. Yanikomeroglu, "On the impact of correlated shadowing on the performance of user-in-the-loop for mobility," in *Proc. UNet-Workshop at the IEEE International Conference on Communications (ICC)*, Ottawa, Canada, June 2012.

[14] H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin, "Diversity in smartphone usage," in *MobiSys '10: Proceedings of the 8th international conference on Mobile systems, applications and services*. New York, NY, USA: ACM, 2010.