Optimal Bandwidth Allocation and QoS-adaptive Control Co-design for Networked Control Systems

Kun Ji and Won-jong Kim*

Abstract: In this paper, we present a co-design methodology of dynamic optimal network-bandwidth allocation (ONBA) and adaptive control for networked control systems (NCSs) to optimize overall control performance and reduce total network-bandwidth usage. The proposed dynamic co-design strategy integrates adaptive feedback control with real-time scheduling. As part of this co-design methodology, a "closed-loop" ONBA algorithm for NCSs with communication constraints is presented. Network-bandwidth is dynamically assigned to each control loop according to the quality of performance (QoP) information of each control loop. As another part of the co-design methodology, a network quality of service (QoS)-adaptive control design approach is also presented. The idea is based on calculating new control values with reference to the network QoS parameters such as time delays and packet losses measured online. Simulation results show that this co-design approach significantly improves overall control performance and utilizes less bandwidth compared to static strategies.

Keywords: Adaptive control, networked control system, optimal bandwidth allocation, quality of service.

1. INTRODUCTION

With the development of computer and communication technologies in the last decades, sensors and actuators can now be equipped with network interfaces, being independent nodes in a real-time control system. This gives rise to an NCS with geographically distributed sensors, actuators, and controllers that communicate via a network [1]. To design an NCS, both its control and communication aspects should be considered because the control performance of the NCS's feedback-control loops is limited by the bandwidth of the communication network. For example, the reduction of the sampling interval improves the control loop's performance [2]. However, a shorter sampling interval requires more network bandwidth to transmit more sensor or control data, which increases the network traffic load. This may affect the system stability and performance of the control loop if the maximum available network bandwidth is exceeded. Therefore, a co-design of control and network-bandwidth allocation must be applied when an NCS is designed [3,4].

Traditionally, research on bandwidth allocation and scheduling techniques focused on static strategies that would ensure average control performance at the expense of permanently occupying the available bandwidth. Hong [5] and Hong and Kim [6] developed a scheduling algorithm to determine the sampling periods of multiple control loops with cyclic service discipline. Thus the performance requirement of each control loop was satisfied, and the utilization of network resources increased. Branicky *et al.* [7] formulated a static optimal scheduling problem under both rate-monotonic-schedulability constraints and NCS-stability constraints.

From the control perspective, the static bandwidth allocation method is an "open-loop" solution in the sense that the static scheduling will not be adjusted at the run time once established at the system set-up. Given sufficient bandwidth, the static bandwidth allocation can successfully guarantee real-time communication and meet the control requirements. However, due to network bandwidth limitation in some applications, not all control loops can simultaneously obtain arbitrarily large bandwidth allocation to provide the best possible control performance. Thus scheduling the network with an "open-loop" algorithm may cause critical messages to fail in timely transmission, degrading control performance, or even leading to instability of certain

Manuscript received June 15, 2007; accepted March 14, 2008. Recommended by Editorial Board member Young Il Lee under the direction of Editor Young-Hoon Joo. This work was supported by the National Science Foundation under Grant No. CMS-0116642.

Kun Ji is with Siemens Technology to Business Center, 1995 University Ave., Berkeley, CA 94704, U.S.A. (e-mail: kunjitamu@gmail.com).

Won-jong Kim is with the Department of Mechanical Engineering, Texas A&M University, College Station, TX 77843-3123, U.S.A. (e-mail: wjkim@tamu.edu).

* Corresponding author.

control loops.

To address these problems, we followed a "closed-loop" philosophy and developed a dynamic ONBA algorithm that makes scheduling decisions based on the performance information of each control loop. Following the methodology of feedback scheduling [8], the control requirements are integrated into the scheduling of the shared network. Bandwidth allocation is implemented as a feedback scheduler in each control loop to assign the optimal sampling frequency to each control loop considering the performance information. This information can be defined based on control-system specifications such as system overshoot, steady-state error, rise time, and so on. The objective of this paper is to present a co-design algorithm for NCS design to maximize the overall quality of control (QoC) through optimally allocating network resources, especially when the available network bandwidth is limited. It can also be used to enable existing NCSs to provide satisfactory QoP under resource constraints.

Dynamic strategies for network scheduling in NCSs can be found in literature. A control-loop scheduling technique called Large Error First (LEF) was presented in [9]. The LEF algorithm used feedback information from the application to assign communication bandwidth to each individual node. However, the relative importance of different control loops in the whole system at different levels was not considered, and the implementation issues of the LEF technique remain to be addressed. In [10] a dynamic arbitration technique called Maximum-Error-First with Try-Once-Discard (MEF-TOD) was presented to grant network access to the control loop with the highest error. However, both LEF and MEF-TOD techniques are referred to communication protocols, and adopting them in existing applications may be very expensive due to the requirements of excessive time and cost for system updates.

A method for optimal off-line scheduling of limited communication resources used for control purpose was presented in [11]. Park *et al.* [12] presented a scheduling method for NCSs to adjust the sampling period as small as possible, allocate the network bandwidth for three types of data, and exchange the transmission orders of data for sensors and actuators. The sampling adjustment was considered for the control analysis, but it was not performed according to the systems dynamics and performance. In [13], bandwidth management of each control loop was done locally at the run-time according to the states of each controlled process, and control laws were designed to account for the variations on the assigned bandwidth. In this approach, however, sampling periods were time-varying, and the quickly varying states might introduce abrupt and too frequent changes from a sampling period to another, which would imply excessive switching between different closed-loop modes (chattering). Further work [14] extended this dynamic bandwidth management to an optimal bandwidth-allocation policy whose complexity might increase the requirement of computational power. Our approach is similar but significantly reduces the computational requirement. In our algorithm, the sampling periods only change among three control modes to be described in Section 3 with little chattering.

The change of the system configuration might also modify the time-delay signature of a networked device, thus change the network QoS. Optimized QoC can be achieved if the networked controller can adapt its control law according to the QoS changes. This can be formulated as an adaptive-control problem that can adjust its parameters on-line according to the changing network QoS. The bandwidth allocation problem and the adaptive control problem are two different problems but follow similar design mechanisms. A bandwidth allocation problem seeks the optimal sampling period for control design based on the control system QoP whereas an adaptive control problem, the optimal control parameters for control design based on the network QoS. Thus they can be formulated as a co-design problem as proposed in this paper. The objective of this co-design problem is to design a networked controller that can adaptively modify the control algorithm according to the control QoP and network QoS.

This paper is organized as follows. The NCS system model and the problem statement are presented in Section 2. The ONBA algorithm is presented in Section 3. QoS-adaptive control design is presented in Section 4. Simulation verification results are given in Section 5.

2. SYSTEM MODEL AND PROBLEM STATEMENT

2.1. ONBA for NCSs

We consider the NCS shown in Fig. 1 with N control loops each controlling a plant. We assume that all the N control loops are independent of each other. The ONBA algorithm is implemented as a feedback scheduler in each control loop to assign the optimal sampling frequency $h_{i,k}$ to each control loop considering the performance, in this specific case, the system error E_i. With the consideration of time-varying sampling frequencies, each plant in closed loop can be described by the following system model.

$$\mathbf{x}_i(k+1) = \Phi(h_{i,k})\mathbf{x}_i(k) + \Gamma(h_{i,k})\mathbf{u}_i(k),\ i = 1,\dots,N, \tag{1}$$

where $\mathbf{x}_i(k) \in R^n$ is the state of Plant i, $\mathbf{u}_i(k) \in R^m$ is the control input for Control Loop i, $h_{i,k}$ denotes the

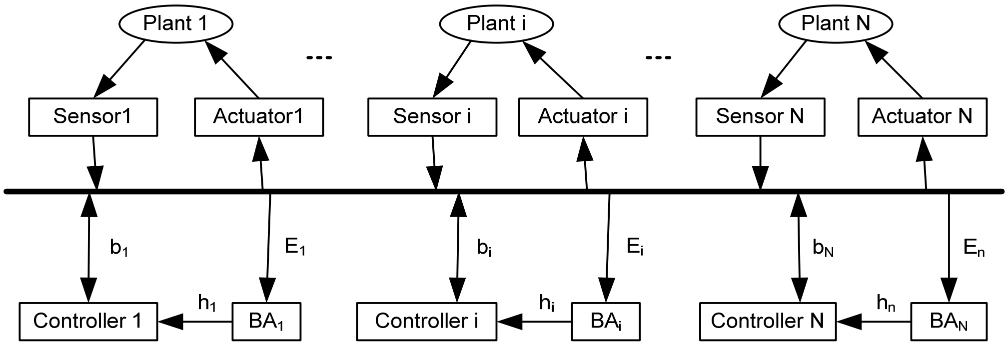


Fig. 1. Networked control systems with ONBA.

sampling period of Control Loop i at time instant k, and $\Phi(h_{i,k})$ and $\Gamma(h_{i,k})$ are real matrix functions of $h_{i,k}$ of appropriate dimensions.

The sampling period $h_{i,k}$ can be obtained from the bandwidth utilization $b_{i,k}$ to be assigned to Control Loop i at time instant k according to the following equation [12].

$$h_{i,k} = \frac{\tau_i}{b_{i,k}}, \quad 0 \le b_{i,k} \le 1, \tag{2}$$

where τ_i denotes the operation time required to finish a control action for Control Loop i in the best case, which only includes the time for data processing such as sampling the sensor, calculating the control output, actuating the actuator, and analog-to-digital and digital-to-analog conversion, and the time for transmitting the data packets from the sensor node to the controller node and from the controller node to the actuator node. The bandwidth utilization $b_{i,k}$ is the parameter which indicates the portion of the network bandwidth assigned to Control Loop i at time instant k. The network utilization is defined to be the ratio of the total time used to transmit data and the total running time [3]. Small network utilization implies that there is much network bandwidth available for other functionalities or control purposes. If the network utilization approaches one, the network becomes saturated, and it is difficult to increase the sampling rates of control devices or add more devices. Then either network bandwidth reallocation to reassign the traffic load or network redesign is needed.

We assume that the periodic sensor data from the sensor node and the control data from the controller node are packetized to an identical bit length L_i. If the data rate of the network medium is R, then the best-case one-way data-transmission time is

$$T_{i,t} = \frac{L_i}{R}. \tag{3}$$

Let $T_{i,p}$ be the time needed for data processing for Control Loop i. The operation time τ_i can be expressed as

$$\tau_i = 2T_{i,t} + T_{i,p}. \tag{4}$$

For each known Control Loop i, $T_{i,p}$ and $T_{i,t}$ in the best case can be measured and computed. Thus from (4) we can assume that τ_i is a constant, then from (2), smaller $h_{i,k}$ indicates bigger $b_{i,k}$. Therefore (2) follows the fact that a control loop with a higher sampling frequency requires more bandwidth allocation to transfer more data. Regarding (2), there are several special cases.

1. When $h_{i,k} = \tau_i$, i.e., $b_{i,k} = 1$, the 100% of the network bandwidth is used by Control Loop i, and no other control loops are allowed to share the network bandwidth. This is the case of NCSs with only a single control loop.
2. When $h_{i,k} = D_i$, where D_i is the maximum allowable loop delay (MALP) for Control Loop i, the minimum network bandwidth utilization of Control Loop i is

$$(b_{i,k})_{\min} = \tau_i / D_i. \tag{5}$$

3. When there are N control loops, the most available bandwidth could be assigned to Control Loop i is (6) while all the other control loops are assumed to use their minimum bandwidths, i.e.,

$$(b_{i,k})_{\max} = 1 - \sum_{j \ne i}^{N} (b_{j,k})_{\min}. \tag{6}$$

In some cases, due to network bandwidth limitation, not all systems can simultaneously obtain enough bandwidth allocation to transfer data and execute at their highest possible sampling frequency. How to perform the optimal bandwidth allocation to obtain the optimal control performance for each closed-loop system is the main research question of this paper. The following rationale is considered: when a controlled plant is in equilibrium, the pre-assigned execution rate (or sampling period) may not be required. That is, the assigned bandwidth can be reduced for the sake of saving overall bandwidth usage and enhancing the bandwidth utilization of other control loops. On the

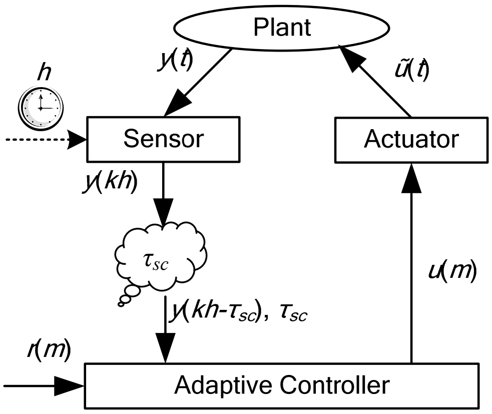


Fig. 2. NCS architecture with one-way time delay.

other hand, when a controlled plant is perturbed, increasing its assigned bandwidth by taking the underutilized bandwidths away from other control loops in equilibrium may hasten system recovery from the perturbation and improve its system performance.

2.2. QoS-adaptive control for NCSs

We consider a control loop in an NCS as shown in Fig. 2. The sensor sampling period is h, k denotes the index of sensor sampling instant, and m denotes the index of control-output ($u(m)$) calculation. The input to the controller is $r(m)$.

The following assumptions are made.

1. Overall time delay is aggregated to be τ_{sc} between the sensor node and the controller node. The delay range is between τ_{min} and τ_{max}.
2. The sensor is clock-driven whereas the controller and the actuator are event-driven. The controller only performs a new calculation after a sensor data-packet has been received.
3. For each sensor data-packet, the time delay τ_{sc} can be measured by assigning a timestamp to each packet. With this timestamp, each data-packet is also numbered.

With these assumptions, the time delay and packet loss information is known at each time instant when sensor data arrive at the controller node. Using this information as the network QoS parameter, a QoS-adaptive controller is proposed in Section 4.

2.3. Control performance analysis of NCSs

Ray and Halevi [15] showed that the feedback-control performance directly depends on the loop delay, which is defined as the interval between the instant when the sensor node samples data and the instant when the actuator actuates the control command. In order to guarantee system stability and adequate control performance, two control measures can be used to determine the maximum allowable loop delay [3]: phase margin ϕ and the closed-loop bandwidth ω_{bw}. To ensure stable and acceptable control performance, the rule of thumb in digital control is that the reasonable sampling rate should be at least 20~40 times as high as the closed-loop bandwidth ω_{bw} [3], i.e., $20 \le \omega_s / \omega_{bw} \le 40$, where ω_s is the sampling frequency. Thus the maximum allowable loop delay D_i for Control Loop i could be estimated by

$$D_i \le T_{i,bw} / 20, \tag{7}$$

where $T_{i,bw} = 2\pi / \omega_{i,bw}$ and $\omega_{i,bw}$ is the closed-loop bandwidth of control loop i.

From another rule of thumb, the sampling period should be 4~10 times as fast as the rise time t_r [2], i.e., $4 \le t_r / T_s \le 10$, where T_s is the sampling period. Thus the maximum allowable loop delay D_i for Control Loop i could also be estimated by

$$D_i \le t_{i,r} / 4, \tag{8}$$

where $t_{i,r}$ is the rise time of closed-loop system i.

Integral of the absolute value of the error (IAE) and integral of the time multiplied by the absolute value of the error (ITAE) are two criteria generally used to evaluate control system performance. They are formulated in continuous time as [2,3]

$$\begin{aligned} IAE &= \int_{t_0}^{t_f} |e|\, dt, \\ ITAE &= \int_{t_0}^{t_f} t|e|\, dt. \end{aligned} \tag{9}$$

In discrete-time as

$$\begin{aligned} IAE &= \sum_{k=k_0}^{k_f} |e_k|, \\ ITAE &= \sum_{k=k_0}^{k_f} k|e_k|, \end{aligned} \tag{10}$$

where t_0 (or k_0) and t_f (or k_f) are the initial and final times of the evaluation period and e is the system error defined as the error between the actual and reference trajectories.

To evaluate the QoC, the following measure is defined as [16]

$$QoC = \frac{1}{IAE}, \tag{11}$$

where IAE is the control performance measure defined in (9)–(10). Equation (11) basically indicates that lower IAE means better QoC. Let $e_{i,k}$ denote the error of Control Loop i at time instant k. It can be expressed in a regulation problem as following

[13,14].

$$e_{i,k} = |x_i(k)| \quad (12)$$

For Control Loop i, we can define a performance criterion that relates control performance (like error) with bandwidth utilization as

$$e_{i,k} = E(b_{i,k}). \quad (13)$$

In general, the less bandwidth allocation is, the worse the control performance (i.e., the larger the error). Thus (13) can be approximated by a linear relation as [12]

$$e_{i,k} = E(b_{i,k}) \approx \frac{\beta_i}{b_{i,k}}, \quad (14)$$

where the parameter β_i is specific to each control loop and can be determined prior to the implementation of the NCSs by evaluating the control performance of each control loop for a broad range of sampling rates or bandwidth allocations. Such linearization method was mentioned in [14].

3. OPTIMAL NETWORK BANDWIDTH ALLOCATION

The ONBA problem is that for a given network with limited bandwidth, how to assign a bandwidth utilization b_i to each control loop according to the control performance and network bandwidth availability such that the overall QoC of the NCSs is optimized. Since each control loop has its own control objective and perturbation situation, only local control optimization is considered here.

The constraint of bandwidth allocation is $\sum_{i=1}^{N} b_{i,k} \le 1$, i.e., the total bandwidth utilization must not exceed the whole network capacity. Then the current additionally available bandwidth utilization is

$$b_a = 1 - \sum_{i=1}^{N} b_{i,k}. \quad (15)$$

If $\sum_{i=1}^{N} (b_{i,k})_{\min} \ge 1$, then the NCSs are not schedulable with the current choice of network. Choosing another network or reducing the number of the control loops is needed.

If each control loop is allocated with a fixed bandwidth, there may be a waste of the network resources in case each control loop's actual bandwidth requirement is less than its fixed bandwidth. In order to provide services to the maximum number of control loops with their QoC requirements and to achieve high utilization of the bandwidth resources, the bandwidth allocated to each control loop needs to be minimized without much degrading its performance.

Thus we formulate the following cost function to be minimized

$$J_{i,k} = a_{i,1} e_{i,k}{}^2 + a_{i,2} b_{i,k}{}^2, \quad (16)$$

where $a_{i,1}$ and $a_{i,2}$ are the weighting coefficients which are selected during engineering time for a tradeoff between local control performance and bandwidth utilization based on (14). The optimization object of bandwidth allocation is to find a suitable bandwidth utilization $b_{i,k}$ that can minimize the overall network bandwidth usage and maximize the system performance (i.e., minimize the error). Minimal instantaneous error leads to minimal integration of the errors based on (9) or (10) and maximal QoC based on (11).

Considering all the control loops, the optimization function J_k for the whole system becomes

$$J_k = \sum_{i=1}^{N} J_{i,k} = \sum_{i=1}^{N} (a_{i,1} e_{i,k}{}^2 + a_{i,2} b_{i,k}{}^2) \quad (17)$$

with the constraint $\sum_{i=1}^{N} b_{i,k} \le 1$. Since all N control loops are independent, J_k is minimal if each $J_{i,k}$ of the ith control loop is minimal. Hence, the optimal bandwidth allocation for each control loop that achieves the optimization of the cost function in (17) can be established.

There can be three notable control modes:

1. When Control Loop i is in equilibrium, i.e., $e_{i,k} \cong 0$, from (5) and (16) we have the optimal bandwidth allocation for Control Loop i as

$$(b_{i,k})_{\text{opt}} = (b_{i,k})_{\min} = \tau_i / D_i, \quad (18)$$

where $(b_{i,k})_{\text{opt}}$ is the optimal bandwidth utilization for Control Loop i at time instant k. Substituting (7) or (8) into (18) yields following (19) or (20).

$$(b_{i,k})_{\text{opt}} = 20\ \tau_i / T_{i,bw} \quad (19)$$

$$(b_{i,k})_{\text{opt}} = 4\ \tau_i / t_{i,r} \quad (20)$$

The corresponding optimal sampling period is

$$(h_{i,k})_{\text{opt}} = T_{i,bw} / 20 \quad (21)$$

$$(h_{i,k})_{\text{opt}} = t_{i,r} / 4. \quad (22)$$

2. When Control Loop i experiences perturbation, $e_{i,k} > e_{th}$, where e_{th} is the preset threshold error. Substituting (14) into (16) and differentiating $J_{i,k}$

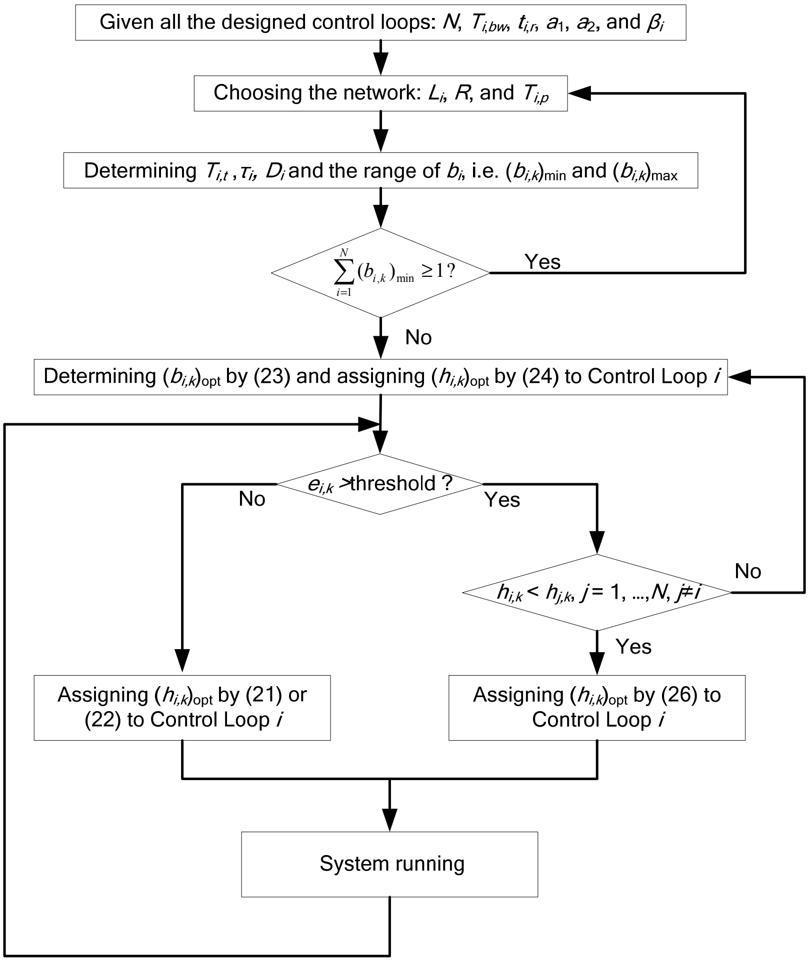


Fig. 3. Optimal bandwidth allocation algorithm.

with respect to $b_{i,k}$, the optimal value $(b_{i,k})_{\text{opt}}$ can be obtained as

$$(b_{i,k})_{\text{opt}} = \sqrt[4]{\frac{a_{i,1}\beta_i^2}{a_{i,2}}}. \tag{23}$$

The corresponding optimal sampling period is with (2)

$$(h_{i,k})_{\text{opt}} = \frac{\tau_i}{\sqrt[4]{\frac{a_{i,1}\beta_i^2}{a_{i,2}}}}. \tag{24}$$

3. If Control Loop i is the only control loop that experiences perturbation or has the highest processing demands, i.e., $h_{i,k} \le h_{j,k}, j = 1, \ldots, N, j \ne i$, all the additional available bandwidth allocation can be assigned to improve its control performance, i.e., $(b_{i,k})_{\text{opt}} = \sqrt[4]{\frac{a_{i,1}\beta_i^2}{a_{i,2}}} + b_a$, where b_a is the current additionally available bandwidth utilization defined in (15). Thus.

$$(b_{i,k})_{\text{opt}} = \sqrt[4]{\frac{a_1\beta_i^2}{a_2}} + 1 - \sum_{i}^{N} b_{i,k} = 1 - \sum_{j \ne i}^{N} b_{j,k} \tag{25}$$

and the optimal sampling period is

$$(h_{i,k})_{\text{opt}} = \frac{\tau_i}{1 - \sum_{j \ne i}^{N} b_{j,k}}. \tag{26}$$

Based on the analysis above, the ONBA algorithm can be summarized in a flow chart given in Fig. 3.

This ONBA can be implemented as a part of the control algorithm in each control loop. For each control loop, the controllers with three different sampling periods according to the three aforementioned control modes are designed prior to the system implementation. During the system run-time, each control loop keeps monitoring the system error to check if it is within a preset threshold, then the decision of which controller should be used is made based on this system error information. If system error is within a preset threshold, then a smaller network bandwidth utilization i.e., a larger sampling period based on (21) or (22) is assigned to this control loop. When the system error is large due to perturbation, there are two cases. If the current control loop is the only control loop that experiences perturbation or it has the highest processing demands, then the optimal bandwidth utilization based on (26) is assigned. i.e., all the currently available bandwidth can be assigned to this control loop to ensure its best QoC. Otherwise, the optimal bandwidth utilization based on (24) is assigned to optimize the overall QoC.

Remark 1: This ONBA technique requires controllers capable of running with different sampling frequencies. For the systems described as (1), controllers are designed with specifying three sampling periods using (21) or (22), (24), and (26) and adapting their gains accordingly, for which the closed-loop stability and performance requirements are met. These three controllers can be designed a priori. Control Mode 3 (26) depends on all control loops, for local optimization, other controller loops are assumed to use Control Mode 2 (23). It is necessary that some parameters of both the individual control loops and the communication networks have to be selected, measured, or identified before running the algorithms, this engineering cost is a common requirement in practical industrial automation and control systems where process identification is always the first step before controller design. Precise knowledge of the type and characteristics of the process to be controlled is indispensable for structuring and designing the controller [17].

Remark 2: This algorithm can be easily extended to cover the case that there are two or more control loops simultaneously experience perturbations by introducing a prioritization mechanism. The decision of assigning additional bandwidth can be made based on the priorities of these control loops according to their processing demands and QoP specifications. At each time, only one control loop can get the additional bandwidth such that the overall bandwidth allocation will not exceed the bandwidth constraint.

4. QOS-ADAPTIVE CONTROLLER DESIGN

For the plant shown in Fig. 2, a continuous-time controller can be designed by standard methodologies before introducing the network. The discrete-time form of the controller can be derived using a continuous-time to discrete-time transformation [2]. The QoS-adaptive control design procedure can be summarized as follows.

1. A simulation is conducted to obtain the n optimal control parameters that maximize the QoC defined in (11) to be the original set of control parameters.
2. In this step, l values of the time delay, τ_l distributed over the range between τ_{min} and τ_{max} are used as constant delays in simulation to search for the optimal control parameters according to different delays. For each τ_l, n optimal control parameters that maximize the QoC can be found by simulation. Then l sets of control parameters can be found and defined as

$$K_{\tau_l} = (k_1, \cdots, k_{n-1}, k_n)_l, \quad l = 1, 2, 3, \cdots. \tag{27}$$

Each K_{τ_l} represents the optimal set of n control parameters in terms of QoC according to a specific constant delay τ_l. This l sets of parameters are then stored in a look-up table in the controller node.
3. The network is introduced between the sensor node and the controller node. At each time instant when the sensor-data packet arrives at the controller node, the time delay τ_{sc} is measured by the controller by checking the timestamp. Based on this measured time delay, an appropriate set of control parameters are selected from K_{τ_l} in (27) stored in a look-up table and used in the current control algorithm. Since there are only l sets of control parameters obtained from the second step, the set of parameters according to the delay closest to the measured delay will be used. If there is no delay, the original set of parameters obtained from the first step will be used.

To deal with the packet losses, varying sampling periods can be used according to the data-packet-loss rate. Under Assumption 3 in Section 3, each data packet is numbered. Let i_k be the packet number, then the varying sampling period used in control algorithm is given by

$$h_m = h(i_k - i_{k-1}), \quad i_k \in \{1, 2, 3, \cdots\}, \tag{28}$$

where h_m is the sampling period used in the mth calculation of the controller, i_k and i_{k-1} are the packet numbers of the current and the last received packet, respectively. If there is no packet loss, h_m is equal to h, and h_m is bounded under Assumption 1 in Section 3.

Remark 3: This QoS-adaptive control method can be extended to cover the situation when time delays and packet losses both occur during the data communication process. This can be done by introducing one more variable representing the sampling period, to the aforementioned look-up table. That is, the optimal set of control parameters is evaluated not only based on the time delay, but also the sampling period. Thus during the run-time, before a new control calculation is performed, the optimal set of control parameters K_{τ_l}, and the new sampling period h_m must be chosen.

Remark 4: A necessary and sufficient stability condition for the system stability with an adaptive controller subject to varying time delays was presented in [16]. Two sufficient conditions were also presented in [16]. The closed-loop stability of the system with QoS-adaptive controller proposed in this section can also be evaluated by using these conditions.

Remark 5: The QoS-adaptive controller design approach developed in this section has a similar design mechanism to the ONBA algorithm proposed in Section 3. They both follow the "real-time feedback" technique to improve the controller design. The ONBA algorithm in a network scheduling problem searches for the optimal sampling period based on control system QoP, whereas the QoS-adaptive control approach searches for the optimal control parameters based on the communication system QoS. This intertwined nature between these two problems requires the co-design of control systems and communication systems in NCS design. The network QoS should be analyzed together with specifying the control QoP before the implementation of the real-time control over networks is completed.

5. SIMULATIONS

In order to verify the ONBA algorithm developed in Section 3, we provide key simulations developed with MATLAB/Simulink.

5.1. Simulation for ONBA

Our illustrative NCS simulation setup contains 5 independent DC motor control loops similar to the configuration shown in Fig. 1. The DC motor system model is given as

$$G(s)=\frac{1}{s(s+1)}. \tag{29}$$

The lead controller is given as

$$D(s)=10\frac{s/2+1}{s/10+1}. \tag{30}$$

All controllers and all DC motors are assumed to be identical which simplifies the performance analysis. We assume the network and control parameters as $\tau = 0.03$ s and $D = 0.3$ s, where τ is the time required to finish a closed-loop control operation, and D is the maximum loop delay as defined in (5). From (5) and (6), we can obtain the working range of bandwidth allocation as

$$\begin{aligned}(b_k)_{\min} &= \tau / D = 0.1,\\ (b_k)_{\max} &= 1 - 4\times(b_k)_{\min} = 0.6.\end{aligned} \tag{31}$$

The sampling periods are $h_{\min} = 0.05$ s and $h_{\max} = 0.3$ s.

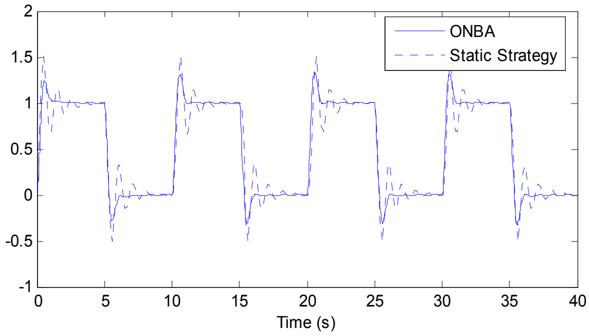


Fig. 4. Step-response comparison.

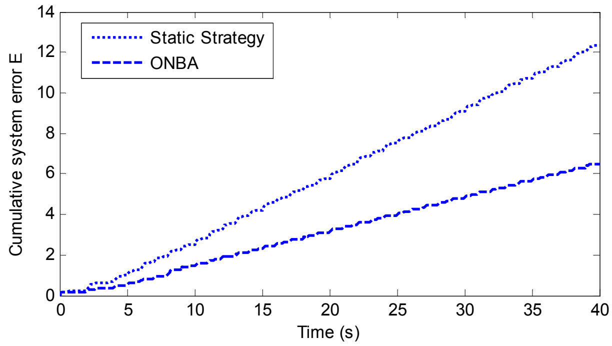
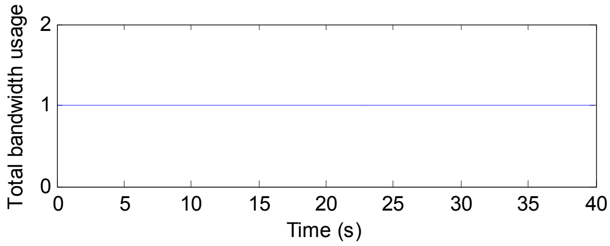


Fig. 5. Cumulative system-error comparison.

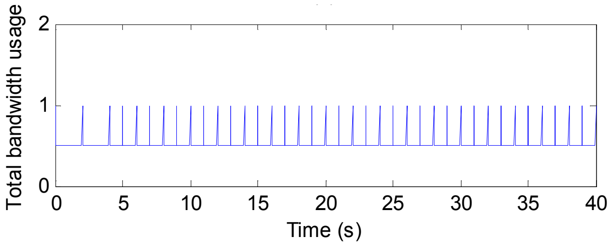
With a static strategy, we would assume all the plants share the bandwidth equally and there is no other traffic load. Thus the bandwidth allocation for each control loop with a static strategy would be $b_k = 1/5 = 0.2$ and the sampling period is $h_k = \tau / b_k = 0.03/0.2 = 0.15$ s.

Both this static bandwidth allocation strategy and the ONBA strategy were implemented to provide a direct comparison between their performances. In the simulation, 5 periodic step disturbances with different phase shifts were inputted to these 5 DC motors, respectively. The system responses of these 5 DC motors are identical except for the phase shift, and the system responses of one of the DC motors with the two strategies are shown in Fig. 4. As evident from Fig. 4, the performance of system with ONBA (in solid line) is better than that with static strategy (in dashed line).

Fig. 5 shows the comparison of the cumulative system errors in these two systems. The closed-loop system error is defined as the absolute difference between the desired response (set-point) and the actual response (feedback output) of the controlled plant. The cumulative system error E is the total cumulative closed-loop system error of all the 5 DC motor control loops, i.e., $E(t)=\int_0^t \sum_{i=1}^{5} |e_i(s)|\, ds.$ We can see that the ONBA strategy achieves better performance than the static strategy by reducing 50% of the cumulative error.



(a) Static strategy.



(b) ONBA.

Fig. 6. Total bandwidth usages comparison.

Fig. 6 shows the comparison of the total bandwidth usages of these two systems. The straight line in Fig. 6(a) indicates that in static strategy, network bandwidth is totally occupied by the 5 DC motor control loops, and no more plants or functionalities could be added. In Fig. 6(b), the network bandwidth usage goes up to 100% only when there are disturbances in the control loops and most of the time the bandwidth occupancy is 50%. Thus some network resource is saved and more control loops or functionalities could be added in this system. From Figs. 4, 5, and 6, it can be concluded that the ONBA algorithm achieves better control performance while using less network bandwidth than the static strategy.

5.2. Simulation for QoS-adaptive control

Another simulation example is presented in this section to demonstrate how the QoS-adaptive controller design proposed in Section 4 can be performed. This example uses plant transfer function taken from [1]. The plant is given by

$$G(s) = \frac{2029.826}{(s+26.29)(s+2.296)}. \tag{32}$$

The PI controller that has been designed in [1] is given by

$$D(s) = \frac{\beta k_P (s + (k_I / k_P))}{s}, \tag{33}$$

where $k_P = 0.1701$, $k_I = 0.378$, and β is a parameter to adjust controller gains. How to choose β to obtain optimal control performance when there is time delay in the control loop is investigated in [1]. If there is no time delay, $\beta = 1$. Using the MATLAB "c2d" command based on Tustin's method, the following discrete-time control law can be derived [1].

$$\begin{aligned} u(i_k h) = {} & u(i_{k-1}h) + [k_p + 0.5k_I h]e(i_k h) \\ & + [0.5k_I h - k_p]e(i_{k-1}h) \end{aligned} \tag{34}$$

Thus in this example, there are two control parameters. (i.e., $n = 2$), $k_1 = \beta k_P$, and $k_2 = \beta k_I$. The sampling period was set to be $h = 30$ ms representing 25% of the rise time of the continuous closed loop based on the "rule of thumb" in (22).

Several simulations were conducted to obtain the optimal sets of control parameters (k_1 and k_2) for different time delays. Table 1 shows the resulting look-up table that relates the optimal sets of control parameters and different time delays.

Then random delays were introduced between the controller node and the sensor node in the DC motor system. The time delays were varying between 10 ms and 100 ms. PI controllers with the control parameters in Table 1 were implemented respectively. The QoS-adaptive controller proposed in Section 4 was also implemented, and the look-up table obtained above (Table 1) was also used. Fig. 7 shows a comparison of the system performances with the adaptive controller

Table 1. Optimal control parameters with various time delays (h = 30 ms).

Time delay (ms)	k_1	k_2	β
0	0.1701	0.3780	1
15	0.1531	0.3402	0.9
30	0.1361	0.3024	0.8
45	0.1106	0.2457	0.65
60	0.0851	0.1890	0.5
75	0.0680	0.1512	0.4
90	0.0501	0.1134	0.3

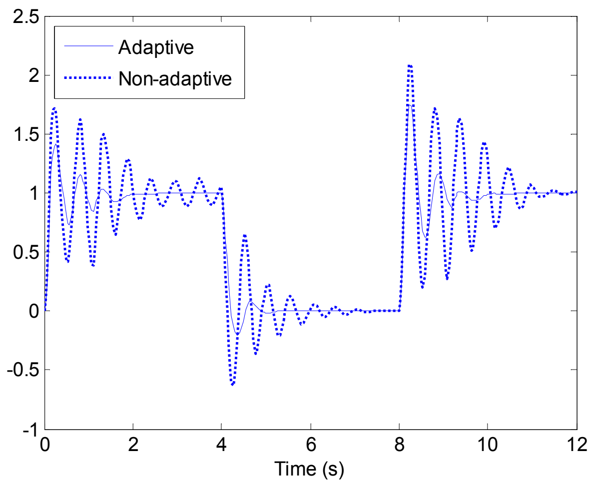


Fig. 7. System responses with the adaptive controller and a non-adaptive (k_1 = 0.1361 and k_2 = 0.3024) controller.

Table 2. QoC values with different control parameters.

Non-Adaptive Controller			QoC
k_1	k_2	β	
0.1701	0.3780	1	0.4520
0.1531	0.3402	0.9	0.6263
0.1361	0.3024	0.8	0.6556
0.1106	0.2457	0.65	0.6405
0.0851	0.1890	0.5	0.6882
0.0680	0.1512	0.4	0.7253
0.0501	0.1134	0.3	0.7714
Adaptive Controller			0.9091

and a non-adaptive controller. Unit pulses with duration of 4 s were used as the input. The solid line in Fig. 7 denotes the system response with adaptive controller, and the dotted line, with non-adaptive controller with fixed control parameters (k_1 = 0.1361 and k_2 = 0.3024). The adaptive controller exhibited better control performance than the non-adaptive controller.

The QoC measure defined in (11) was used to further verify the effectiveness of the proposed adaptive controller. Simulations of the DC motor system with various sets of control parameters shown in Table 1 were conducted and corresponding QoC values are calculated and shown in Table 2. The QoC value of system with the QoS-adaptive controller is also shown in this table to have a direct comparison with these non-adaptive controllers. The QoC value with the QoS-adaptive controller is better than those with any of the non-adaptive controllers given in Table 1. Thus Table 2 proves that QoC improvement can be achieved by implementing the QoS-adaptive controller proposed in Section 4. The simulation results in Sections 5.1 and 5.2 also verified that the system stability is not affected by the control design methodologies proposed in Sections 3 and 4.

6. CONCLUSIONS

In this paper, novel co-design of network bandwidth allocation and the QoS-adaptive control strategy was proposed. As part of this co-design methodology, a dynamic optimal network bandwidth allocation algorithm for NCSs with communication constraints was developed. The proposed dynamic strategy integrates feedback control with real-time network scheduling and it makes scheduling decisions based on the dynamic QoP information of each control loop.

The foremost advantage of this algorithm is that the computational-power requirement is low and that the allocation of bandwidth to control loops can be done locally at run time according to how far the control loops are from their equilibriums. Thus it can be implemented as a feedback scheduler in existing control applications and reduces system engineering cost. Another key advantage is that this algorithm does not cause excessive switching between different closed-loop modes (chattering) which may lead to instability. The simulation results showed that this ONBA approach better utilized the bandwidth resource under the constraint of the limited available bandwidth in comparison with the static approach.

As another ingredient of the co-design methodology, a QoS-adaptive control design methodology was also presented. This methodology is based on calculating new control values with reference to the QoS parameters such as time delays and packet losses measured online in real time. The simulation results demonstrated that the QoC could be improved by implementing the QoS-adaptive controller proposed in this paper.

REFERENCES

[1] Y. Tipsuwan and M.-Y. Chow, "Control methodologies in networked control systems," *Control Engineering Practice*, vol. 11, no. 10, pp. 1099-1111, Feb. 2003.

[2] G. F. Franklin, J. D. Powell, and M. L. Workman, *Digital Control of Dynamic Systems*, 5th ed. Reading, Addison-Wesley, MA, 2006.

[3] F.-L. Lian, J. Moyne, and D. Tilbury, "Network design consideration for distributed control systems," *IEEE Trans. on Control Systems Technology*, vol. 10, no. 2, pp. 297-306, Mar. 2002.

[4] K. Ji and W. -J. Kim, "Real-time control of networked control systems via Ethernet," *International Journal of Control, Automation, and Systems*, vol. 3, no. 4, pp. 591-600, Dec. 2005.

[5] S. H. Hong, "Scheduling algorithm of data sampling times in the integrated communication and control systems," *IEEE Trans. on Control Systems Technology*, vol. 3, no. 6, pp. 225-230, June 1995.

[6] S. H. Hong and Y. C. Kim, "Implementation of a bandwidth allocation scheme in a token-passing fieldbus network," *IEEE Trans. on Instrument and Measurement*, vol. 51, no. 2, pp. 246-251, Apr. 2002.

[7] M. S. Branicky, S. M. Phillips, and W. Zhang "Scheduling and feedback co-design for networked control systems," *Proc. of IEEE Conference on Decision and Control*, Las Vegas, pp. 1211-1217, Dec. 2002.

[8] A. Cervin, J. Eker, B. Bernhardsson, and K.-E. Årzén, "Feedback-feedforward scheduling of control tasks," *Journal of Real-Time Systems*,

vol. 23, no. 2, pp. 25-53, 2002.

[9] J. Yepez, P. Marti, and J. M. Fuertes, "Control loop scheduling paradigm in distributed control systems," *Proc. of the 29th IECON*, Roanoke, USA, Nov. 2003.

[10] G. C. Walsh and H. Ye, "Scheduling of networked control systems," *IEEE Control Systems Magazine*, vol. 21, no. 1, pp. 57-65, Feb. 2001.

[11] H. Rehbinder and M. Sanfridson. "Scheduling of a limited communication channel for optimal control," *Automatica*, vol. 40, no. 3, pp. 491-500, Mar. 2004.

[12] H. S. Park, Y. H. Kim, D.-S. Kim, and W. H. Kwon, "A scheduling method for network-based control systems," *IEEE Trans. on Control Systems Technology*, vol. 10, no. 3, pp. 318-330, May 2002.

[13] M. Velasco, J. M. Fuertes, C. Lin, P. Martí, and S. Brandt, "A control approach to bandwidth management in networked control systems," *Proc. of the 30th Annual Conference of the IEEE Industrial Electronics Society*, pp. 2343-2348, Busan, Korea, Nov. 2004.

[14] M. Velasco, P. Martí, and M. Frigola, "Bandwidth management for distributed control of hightly articulated robots," *Proc. of IEEE International Conference on Robotics and Automation*, Barcelona, Spain, pp. 266-271, Apr. 2005.

[15] A. Ray and Y. Halevi, "Integrated communication and control systems: part I-analysis and part II-design consideration," *ASME Journal of Dynamic Systems Measurement and Control*, vol. 110, no. 4, pp. 367-381, Dec. 1988.

[16] P. Marti, J. Yepez, M.Velasco, R. Villa, and J. M. Fuertes, "Managing quality-of-control in network-based control systems by controller and message scheduling co-design," *IEEE Trans. on Industrial Electronics*, vol. 51, no. 6, pp. 1159-1167, Dec. 2004.

[17] "SIMATIC Standard PID Control Manual," Siemens AG, Edition Mar. 2004.

Kun Ji received the B.S. and M.S. degrees in Precision Instrumentation and Engineering from Tsinghua University, Beijing, China, in 1999 and 2002, respectively, and the Ph.D. degree in Mechanical Engineering from Texas A&M University, College Station, in 2006. He has been with Siemens Technology to Business Center, Berkeley, where currently he is a Research Scientist. His research interests focus on analysis, design of networked control systems, real-time control systems, and precision mechatronics systems.

Won-jong Kim received the B.S. (*summa cum laude*) and M.S. degrees in Control and Instrumentation Engineering from Seoul National University, Seoul, Korea, in 1989 and 1991, respectively, and the Ph.D. degree in Electrical Engineering and Computer Science from Massachusetts Institute of Technology (MIT), Cambridge, in 1997.

Since 2000, he has been with the Department of Mechanical Engineering, Texas A&M University (TAMU), College Station, where currently he is an Associate Professor and the Holder of the Dietz Career Development Professorship II. After the Ph.D. degree, he was with SatCon Technology Corporation, Cambridge, MA, for three years. His current research interests include the analysis, design, and real-time control of mechatronic systems, networked control systems, and nanoscale engineering and technology. He is the holder of three U.S. patents on precision positioning systems.

Dr. Kim is a Member of Pi Tau Sigma and American Society of Mechanical Engineers (ASME) and a Senior Member of Institute of Electrical and Electronics Engineers (IEEE). He is a Primary Member of the ASME Mechatronics Technical Committee and a Member of the IEEE Nanotechnology Council. He was the Recipient of Korean Institute of Electrical Engineers' Student Paper Contest grand prize in 1988, Samsung Electronics' Humantech Thesis gold prize for his MIT dissertation in 1997, the NASA Space Act Award in 2002, and the 2005 Professional Engineering Publishing Award for the best paper published in 2004 in *Journal of Engineering Manufacture*. He was also a Semifinalist of the National Institute of Standards and Technology (NIST)'s Advanced Technology Program competition in 2000. He was appointed a Select Young Faculty Fellow by TAMU College of Engineering and the Texas Engineering Experiment Station twice in 2003 and 2005. He received the BP Teaching Excellence Award by TAMU College of Engineering in 2006.