2004 IEEE International Conference on Systems, Man and Cybernetics

Survey on the Performance Analysis of Networked Control Systems*

Yu Jianyong **Yu Shimin**
Institute of Intelligent Information System
Zhejiang University of Technology
Hangzhou, 310032, China
yujyong@126.com ysm@zjut.edu.cn

Wang Haiqing
National Laboratory of Industrial Control technology
Zhejiang University
Hangzhou, 310027, China
hqwang@iipc.zju.edu.cn

Abstract: *The insertion of the communication network in the feedback control loop makes the analysis and design of a networked control system more complex, and induces some issues that degrade the control system's performance and even cause system instability. This paper focuses on main aspects around performance analysis of NCSs: network-induced delays, sampling period, jitter, data packet dropout, network scheduling and stability. These issues must be considered in the design of an NCS. Therefore, this work summarizes many research results, and remarks some related handling approaches and techniques. The main purpose of the survey is to present new research state of NCSs and point out some fields of future researches.*

Key words: network-induced delays, jitter, packet dropout, network scheduling, stability

1. Introduction

Feedback control systems wherein the control loops are closed through a real-time network are called networked control systems (NCSs) [1,2,3]. NCSs can deal with all the continuous, discrete and hybrid control asynchronous processes, and support various topologies, including bus, star and tree, which are more flat and stable than the structure used in hierarchical control system. As an alternative to traditional point-to-point communication, the common-bus network architecture of NCSs offer more efficient reconfiguration, better resource utilization, and also reduce installation and maintenance cost. However, the change introduces different forms of time delays uncertainty between sensors, actuators, and controllers. It is well known in control systems that time delays can degrade a system's performance and even cause system instability.

Because of the variability of network-induced time delays, NCSs may be time-varying systems, making analysis and design more challenging. It is natural to analyze an NCS from the discrete-time point of view. For discrete-time models, most researchers assume that the network is synchronized and the sampling rates of sensors, controllers, and actuators are the same. Some work adopted continuous-time models of NCSs, such as [2,4,5,6]. Current research shows that analyzing methods used for an NCS include stochastic Lyapunov function, augmented state space, jump linear systems, and limited communication.

Without loss of generality, a closed-loop block diagram of an NCS is shown in Fig1. Where, the controller node and the actuator node are event-driven while the sensor node is clock-driven.

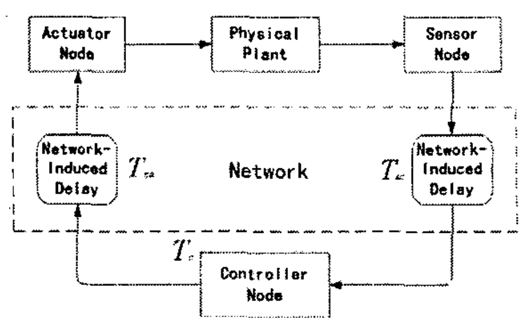

Figure 1: The closed-loop block diagram of an NCS

The remainder of this paper is organized as follows: Sections 2-7 analyze six fundamental issues that influence performance of an NCS respectively: network-induced delays, sampling period, jitter, data packet dropout, network scheduling and stability. Many research results are given and related handling approaches are also remarked. The conclusion and future work are presented in Section 8.

* This work was supported by the National Nature Science Foundation of P. R. China (No.20206028).

2. Network-Induced Delays

In an NCS, various delays with variable length occur due to sharing a common network medium, which are called network-induced delays [7].

Tow classes of time delays are included in an NCS: (1) the delay from a sensor to a corresponding controller, T_{sc}; (2) the delay from a controller to a corresponding actuator, T_{ca}. Network-induced delays may vary widely according to the transmission time of messages and the overhead time. The transmission time through the media is largely dependent on the network protocols, especially data link layer protocols of networks and message length, whereas the overhead time is largely dependent on the network scheduling method.

In order to analyze control systems with network delays in a loop, [8□9] used three models of the network delays: (1) Constant delay; (2) Random delays which are independent between T_{sc} and T_{ca}; (3) Random delay, with probability distributions governed by an underlying Markov chain. Model (1) is the simplest model for all transfers in the communication network and may be a good model even if the network has varying delays. One way to achieve constant delays is by introduction of timed buffers after each transfer. By making these buffers longer than the worst case delay time the transfer time can be seen as being constant. This method to make the communication delays constant was proposed in [10]. A drawback with this method is that the delay time often is longer than necessary, which can lead to decreased performance as shown in [8,9,11]. Network delays are usually random because of several sources, for instance, waiting for the network to become idle; a retransmission of urgent data is needed when transmission errors occur; nodes' waiting for a random time to avoid a collision at the next try when some networks collisions occur, and so on. As the activities in the system usually are not synchronized with each other, the above listed delays will be random. Model (2) assumes the transfer delay is independent of the previous delay and has different probability distributions for T_{sc} and T_{ca}. To model network queues and varying network loads, Model (3) needs to have a memory. One way to model dependence between samples is by letting the distribution of the network delays be governed by the state of an underlying Markov chain. The transitions between different network load states in the communication network are modeled with a Markov chain. The model is closely related to the models used in jump systems, but it's each state postulates different probability distributions for T_{sc} and T_{ca}.

3. Sampling Period

The performance of an NCS is highly dependent on the network sampling time. Increased network sampling time can improve the performance of the system. However, beyond a critical point, sampling frequency begins to adversely affect NCSs performance, because of network loading when the number of messages are close to the network saturation limit, message delivery time delays increase and in some cases become unbounded [12]. Optimizing the performance of an NCS can be achieved by balancing the increasing network sampling frequency with the resulting network performance degradation.

While a shorter sampling period is preferable in most control systems, for some purposes it can be lengthened up to a certain bound within which stability of the system is guaranteed in spite of the performance degradation. This certain bound is called a maximum allowable delay bound (MADB) [7,13,14]. A basic sampling period consists of sampling delay, transmission time of periodic data, transmission time of sporadic data, and transmission time of messages. The largest sampling period in an NCS depends on the largest MADB. [15] presented a sampling period determination algorithm.

In an NCS, a sampling period should be long enough to guarantee real-time transmission of sporadic data and periodic data, and minimum network utilization for a periodic data. However, the sampling period should be short in order to be within the MADB to guarantee the stability of the given system. Sometimes the sampling period may exceed the MADB because of network-induced delays. Therefore, it is important to decrease the sampling period by minimizing network-induced delays.

Due to the interaction of the network and control requirements, the selection of the best sampling period is a compromise. Smaller sampling periods guarantee a better control quality, but result in high frequency communication and may degrade the network quality. The degradation of network quality could further worsen the control quality due to longer time delays when the network traffic is nearly saturated [12,16]. In general when the number of messages on the network increases, system nodes experience longer delays for sending and receiving information [17]. The selection of the best sampling period is presented in [7,12,16,18].

4. Jitter

Jitter is defined by IEEE as "time-related, abrupt, spurious (false) variations in the duration of any specified related interval", and arises due to clock drift, branching in the code, scheduling, communication, and use of certain computer hardware structure, e.g. cache memory. From the view of control, jitter is categorized as: control period jitter, delay jitter and sampling jitter. From the view of scheduling, jitter is categorized as: input jitter, output jitter, queuing jitter and deadline jitter. [19] considered six different jitters, and reduced to three cases: sampling jitter, sampling-actuation delays, and both combined. All kinds of jitter should be as small as possible to improve control performance.

It is noted that vacant sampling and sample rejection are harmful to all computer systems, not only control systems. Jitter degrades performance and causes instability even if for the case when vacant sampling or sample rejection does not occur. The jitter distorts a signal in a control system. The degradation depends very much on the dynamics of the process and on the type of jitter. The only way to get rid of delay jitter is to use a buffer, trading delay for jitter. Stability in case of jitter has been treated by [20] and others. A stochastic model for the jitter is usually assigned for the analysis and design of controllers, e.g. a rectangular distribution around a mean value, or a Markov chain.

The jitter compensation approach in a real-time system needs to be investigated with respect to its implementation cost (computational overhead and memory requirements) and the availability of the necessary information to recalculate the controller parameters when it is needed (information availability). With respect to the implementation cost, at each control task instance execution, the controller parameters must be updated according to the actual jitters. Two strategies may be applied: runtime or offline calculations. If the controller parameter adjustment is performed by online extra calculations according to actual jitters, the introduced computational overhead will depend on the control design method and controller design strategy that is being used. If the computational overhead is not negligible, the controller parameter adjustment can be performed online by accessing offline pre-calculated lookup tables. These tables will contain the necessary parameters to allow the control computation to compensate for sampling jitter and the sampling-actuation delays that may appear at runtime. Some recent research has focused on the jitter problem itself using specific scheduling-based solutions [21, 22,23]. However, even after modifying the scheduling algorithm, in these approaches, jitter is not completely eliminated. [19] proposed to accept the jitter that the scheduling algorithm is bound to introduce and to compensate for it at runtime in the controller design so as to minimize the system degradation that would otherwise occur.

In [24], control delay jitter is investigated by simulation of a servo. A static delay-compensating controller is designed to compensate for random delays, which are close to the expected mean value. [3] also treats control period jitter, timeout and vacant sampling. [19] proposed a new approach for real-time scheduling of control systems by compensating for sampling jitter and sampling-actuation delays through the adjustment of controller parameters.

5. Data Packet Dropout

The network can be viewed as a web of unreliable data transmission paths. Some packets not only suffer from transmission delay but also may be lost during transmission for the worse case. Thus, how such packet dropouts affect the performance of an NCS is an issue that must be considered.

Network packet drops occasionally happen in NCSs when there are node failures or message collisions. Although most network protocols are equipped with transmission-retry mechanisms, they can only retransmit for a limited time. After this time has expired, the packets are dropped. Furthermore, for real-time feedback control data such as sensor measurements and calculated control signals, it may be advantageous to discard the old, un-transmitted message and transmit a new packet if it becomes available. In this way, the controller always receives fresh data for control calculation. Normally, feedback-controlled plants can tolerate a certain amount of data loss, but it is valuable to determine whether the system is stable when only transmitting the packets at a certain rate and to compute acceptable lower bounds on the packet transmission rate.

In [1,15], an NCS with dropouts is modeled as asynchronous switched system. The approach replaces the true-switched system with an "averaged system" and then provides some sufficient stability conditions on the system. Because only average dropout rates are considered, the achieved results may be very conservative. Another important contribution is found in [3]. In this work, the dropouts are modeled by a Markov chain with two states and are treated as vacant sampling. This work proposes two approaches for handling data dropouts: using past control signals or estimating the lost data and computing new control signals. The stability of an optimal LQ controller under the two approaches is analyzed. This work, unfortunately, does not provide a rigorous analysis of the dropout model and only demonstrates the results through examples. [25] extends the method proposed in [26] to dropout processes governed by Markov chains, and assumes that the feedback measurements are randomly dropped with a distribution selected from an underlying Markov chain. The main result of [25] is an equation that expresses the power in the networked control system's output signal as a function of the Markov chain's probability transition matrix.

Network saturation occurs when the network traffic increases beyond a point when not all messages can be delivered to their destinations in time. Messages may experience long delays due to queuing in buffers, or may be lost due to collisions. Network saturation can be retarded by better utilizing network bandwidth. For a given sampling frequency, implementing estimation methods in an NCS would reduce network traffic increasing the effective bandwidth of the system. By defining dead bands on broadcasting nodes [27], the amount of traffic can be reduced

6. Network Scheduling

In feedback control systems, it is necessary to find the maximum allowable delay bound (MADB) for stability of NCSs [7,13], and then to find an appropriate network scheduling method that limits the network-induced delay to less than the MADB. A network scheduling method is required to reduce a basic sampling time within the MADB, while guaranteeing real-time transmission of sporadic and periodic data, and to minimize network utilization for non-real time message. The network in an NCS should handle network scheduling algorithms differing in some characteristics from processor scheduling algorithms, such as the rate monotone scheduling algorithms and the deadline monotone scheduling algorithms. These processor-scheduling algorithms have limitations when applied to NCSs, because a retransmission of periodic data with old values suspended by other urgent data transmissions is meaningless. While the purpose of conventional scheduling methods focused on whether all types of data can be transmitted within a given bandwidth or sampling period, the presented methods focus on that of the conventional methods using MADB as well as setting of transmission order of data. By setting the transmission order of data, the smaller sampling period can be obtained than one from the conventional algorithm.

A scheduling algorithm that can allocate the bandwidth of a network and determine sensor data sampling periods was presented by [28]. In [28], the control system had only single input and single output (SISO), only periodic data were considered, and the MADB was not obtained analytically. A network scheduling method considering three types of data based on a multi-input and multi-output (MIMO) system was proposed by [29]. However, the estimation of MADB using the Ricatti equation is too conservative, which means the estimated MADB is too small, but the network scheduling method discussed in [29] is somewhat heuristic. In [1,2,30,31,32], calculation methods of MADB for stability analysis of an NCS were also presented.

However, these results about calculation of MADB are conservative to be of practical use and still need to be improved. Further research is needed with regard to an estimation of a less conservative MADB for stability of the NCSs and systematic scheduling methods for three types of data are demanded for NCSs. [7] proposed a new method to obtain the MADB guaranteeing stability in terms of linear matrix inequalities (LMI) based on [13,14]. This method gave a much less conservative delay bound than the existing methods. It allocates the bandwidth of a network to a node, determines the sensor data sampling periods of each loop using the obtained MADB, guarantees real-time transmission of sporadic data and periodic data within the sampling periods, and minimizes the network utilization for message. The proposed approach differs from the popular queuing analysis because there is no queue for time-critical periodic data.

7. Stability of NCSs

The stability is a basic problem in the design of an NCS.

The occurrence of transmission events on the network is unknown and often modeled as a random process, e.g., Poisson process, and the resulting interval between each access to the network are independent and have an exponential distribution. The stochastic Lyapunov function method holds much promise for determining almost-sure stability and control system performance. The approach in [31] provided guarantees by employing transmission deadlines. For the first time, [31] proposed a novel protocol, try-once-discard (TOD) protocol, which employs dynamic scheduling, allocating network resource based on the need. In TOD, the node with the greatest weighted error from the last reported value will win the competition for the network resource. Such a method is vulnerable to noise.

The augmented state space method and jump linear control system method are two significant methods proposed in the literature for analyzing and designing an NCS. The former one reduced the problem to a finite dimensional discrete-time control by augmenting the system model to include past values of plant input and output (i.e., delayed variables) as additional states [5,33]. A necessary and sufficient condition for system stability was established only for the special case of periodic delays. This technique is very useful for developing control laws to improve the performance of an NCS [34,35,36] except that it fails to give a general stability condition for random delay. In [37], distributed linear feedback control systems with random communication delays were modeled as a jump linear control systems, in which random variation of system delays corresponds to randomly varying structure of the state-space representation. Necessary and sufficient conditions were found for zero-state mean-square exponential stability of the considered class of systems. This method requires that the transition probability matrix is known a priori. Furthermore, both methods were limited to the one packet transmission problem.

An NCS with data packet dropout can be modeled as an asynchronous dynamical system (ADS) with rate constraints on events. The stability of this type of system is studied in [15,38]. In [1,15], stability of an NCS was also characterized using a hybrid system stability analysis technique and modeled an NCS with packet dropout and multiple-packet transmission (which may occur due to the limitation of the control network) as an asynchronous dynamical system. Stability regions and stability of an NCS have been proposed using a hybrid system technique in [1]. The influence of sampling period to the stability of an NCS has been presented in [2].

8. Conclusions and Future Work

The improvement of NCSs performance can be divided into two areas. First, to further guarantee the determinism of transmission time and reduce the end-to-end time delays, device-processing times should be minimized and network protocols can be improved. Second, advanced optimal or robust controller design can overcome the uncertainty in an NCS and achieve the best control performance.

As the nodes distributed independently, the multi-rate sampling is natural for NCSs, but it always brings about some problems: e.g., the constraint of network bandwidth requests better signal quality; redundant signals cause delay, vacant sampling. It is common to adopt the event-triggered sampling to deal with the constraint of network bandwidth and the negative impact of redundant signals on the performance of system.

While using Markov chains to model an NCS, in fact, the states of Markov chains are usually unknown. The key issue is how to identify the number of the states of Markov chains and their transient probability using HMM (Hidden Markov Model) [39] while designing and analyzing an NCS.

There are several interesting problem still to be solved for future work:

(1). The error bounds of the scheduler for NCSs will be interesting topics. The analysis of worst-case error bound of the scheduler is a design guideline of scheduling method in NCSs.

(2). Though several papers discuss about a stability analysis of NCSs with packet dropout [15,32,40], more problem formulations and scheduling analysis of packet dropping in NCSs will be interesting topics.

(3). The Ethernet technology allows field device like sensors, actuators, and controllers to be interconnected at low cost using less wiring and requiring less maintenance than point-to-point interconnections. However, nondeterministic property of Ethernet protocol is a major problem when Ethernet is applied in control network. Consequently, using switched Ethernet by subdividing the network architecture is another way to overcoming these problems.

References:

[1] W. Zhang, M.S. Branicky, and S.M. Philips, Stability of networked control systems, *IEEE Control Systems Magazine*, 21(1): 84-99, February 2001.

[2] G.C. Walsh, H. Ye, and L. Bushnell, Stability analysis of networked control systems, in *Proceedings of the American Control Conference*, pp. 2876-2880, San Diego, June 1999.

[3] J. Nilsson, Real-Time Control Systems with Delays, Ph.D. dissertation, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden, 1998.

[4] G.C. Walsh, O. Beldiman and L.G. Bushnell. Asymptotic behavior of networked control systems. In *Proceedings of the Conference on Control Applications*, pp1448-1453, Aug.1999.

[5] Yoram, H. and Asok, R. Integrated Communication and Control Systems: Part I - Analysis, *Journal of Dynamic Systems, Measurement, and Control,* 110, Dec., pp.367-373. 1988

[6] Y.H.Kim, W.H.Kwon, and H.S.Park. Stability and a scheduling method for networked-based control systems. In *IECON Proceedings*, volume 2, pp934-939, Aug.1996.

[7] Dong-Sung Kim. A Scheduling Method for Networked Control Systems in Fieldbus Environments. Ph.D Dissertation, Seoul National University, Feburary 2003

[8] Nilson, J., Bermhardsson, B. and Wittermark, Schtochastic Analysis of control of real time systems with random time delays, *Automatica*, 34, (1), pp.57-64. B. 1998.

[9] Nilson,J., Real time control systems with delays, 1998 Ph.D dissertation, Dept. of Automatic control, Lund Institute of Technology, Lund, Sweden.

[10] LUCK, R. and A. RAY . An observer-based compensator for distributed delays. *Automatica*, 26:5, pp: 903-908. 1990.

[11] Nilson, J. Analysis and Design of Real-Time Systems with Random Delays. Lic Tech thesis TFRT-3215, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden. 1996

[12] F.-L. Lian. Analysis, Design, Modeling, and Control of Networked Control Systems. Ph.D thesis, University of Michigan, May 2001.

[13] D.S.Kim, Y.S.Lee, W.H.Kwon, H.S.Park, Maximum Allowable Delay Bounds in Networked Control Systems, *International Journal of Control Engineering Practice*, 2003.

[14] D.S.Kim, Y.S.Lee and W.H.Kwon "Scheduling Analysis of Networked Control system", *IFAC Conference on New Technology on Computer Control*, pp. 77-82, November, 2001

[15] W. Zhang. Stability analysis of networked control systems. PhD Thesis, Case Western Reserve University, 2001.

[16] Feng-Li Lian; Moyne, J.; Tilbury, Network design consideration for distributed control systems . *IEEE Transaction on Control System Technology*. 10, (2), pp: 297-307. March 2002

[17] B. Englert, L. Rudolph, and A. Shvartsman. Developing and refining an adaptive tokenpassing strategy. *21st International Conference on Distributed Computing Systems, 2001*, April 2001.

[18] Feng-Li Lian, James Moyne, Dawn Tilbury, Time Delay Modeling and Sample Time Selection For Networked Control Systems, *Proceedings of ASME-DSC*, New York, November 11-16, 2001.

[19] Pau Martí, Gerhard Fohler, Krithi Ramamritham and Josep M. Fuertes, Jitter Compensation for Real-Time Control Systems, *22nd IEEE Real-Time Systems Symposium*, London, UK, December, 2001

[20] K. Hirai, Y. Satoh, "Stability of a System with Variable Time Delay", *Trans. Automatic Control*, AC-25, p552-554, 1980.

[21] P. Albertos, A. Crespo, I. Ripoll, M. Vallés and P. Balbastre "RT control scheduling to reduce control performance degrading". *39th IEEE Conference on Decision and Control.* Sydney (Australia), December 12-15, 2000.

[22] S. Baruah, G. Buttazzo, S. Gorinsky and G. Lipari, "Scheduling Periodic Tasks Systems to Minimize Output Jitter" *Proc. International Conference on Real-Time Computing Systems and Applications, IEEE Computer Society Press.* pp 62-69, Hong Kong, December 1999.

[23] A. Cervin "Improved Scheduling of Control Tasks" in *Proceedings of the 11th Euromicro Conference on Real-Time Systems*, York, England, June 1999.
[24] N. Andreff, "Robustness to jitter in real-time systems", *ISSN 0280-5316, ISRN LUTFD2/ TFRT--5507--SE*, June 1994.pp.3015-3020.2002
[25] Q. Ling and M.D. Lemmon, Soft real-time scheduling of networked control systems with dropouts governed by a Markov chain, *American Control Conference*, June 2003.
[26] Q. Ling and M.D. Lemmon, "Robust Performance of Soft Real-time Control Systems with Data Dropouts", *IEEE Conference on Decision and Control*, 2002, Las Vegas.
[27] Otanez,P.G., Moyne, J.R. and Tilbury, D.M., 2002, Using deadbands to reduce communication in networked control systems, *Proceedings of the American Control Conference*, 4, pp.3015-3020.
[28] Hong, S.H. Scheduling Algorithm of Data Sampling Times in the Integrated Communication and Control Systems, *IEEE Transaction on Control System Technology*, 3, (2), pp.225-231. Jun 1995
[29] Park, H.S., Kim, Y.H., Kim, D.S and Kwon, W.H. A Scheduling Method for Network Based Control Systems, *IEEE Transaction on Control System Technology*, 10, (3), pp.318-330. May 2002
[30] Branicky, M.S., Phillips, S.R. and Wei Zhang, 2000, Stability of networked control systems: explicit analysis of delay, *Proceedings of the American Control Conference*, 4, pp.2352-2357.
[31] Walsh, G.C., Hong Ye. and Bushnell, L.G. , 2002, Stability analysis of networked control systems, *IEEE Transactions on Control Systems Technology*, 10, (3), pp.438-446.
[32] Walsh, G.C. and Hong Ye. 2001. Scheduling of networked control systems, *IEEE Control Systems Magazine*, 21, (1), pp.57-65.
[33] A. Ray and Y. Halevi, Integrated communication and control systems: Part□-Design considerations, in *ASME Journal of Dynamic Systems, Measurement and Control*, 110: 374-381, December 1988.
[34] L.W. Liou and A. Ray, "A stochastic regulator for integrated communication and control systems: Part I—Formulation of control law," *ASME J. Dynamic Syst., Measurement, Contr.*, vol. 113, pp. 604–611, Dec. 1991.
[35] A. Ray, "Output feedback control under randomly varying distributed delays," *J. Guidance, Contr., Dynamics*, vol. 17, no. 4, pp. 701–711, July–Aug. 1994.
[36] A. Ray, L. W. Liou, and J. H. Shen, "State estimation using randomly delayed measurements," *ASME J. Dynamic Syst., Measurement, Contr.*,vol. 115, pp. 19–26, Mar. 1993.
[37] R. Krtolica and U. Özgüner *et al.*, "Stability of linear feedback systems with random communication delays," in *Proc. Amer. Contr. Conf.*, Boston, MA, June 1991.
[38] A. Hassibi, S.P. Boyd, and J.P. How, "Control of asynchronous dynamical systems with rate constraints on events," in *Proc. IEEE Conf. Decision and Control*, Phoenix, AZ, Dec. 1999, pp. 1345-1351.
[39] H.Y.Kim, C.M.Lee and H.S.Park, Optimal Period and Priority Assignment using Task and Message-based Scheduing in Distributed Control Systems, *Jornal of Control, Automation and System Engineering*, 8, (6), pp.506-513. 2002
[40] Rabello, A., Bhaya, A., Stability of asynchronous dynamical systems with rate constraints and applications, *Proceedings of the American Control Conference*, 2, pp.1284-1289. 2002