Control Engineering Practice 60 (2017) 1–6

Contents lists available at ScienceDirect

Control Engineering Practice

journal homepage: www.elsevier.com/locate/conengprac

A method for setpoint alarming using a normalized index

Timothy I. Salsbury*, Carlos F. Alcala

Controls Research Group at Johnson Controls, Inc. (JCI), 507 E Michigan Street, Milwaukee, WI 53202, USA

ARTICLE INFO

Keywords:
Control performance
PID control
Statistical process control
Normalized index
Big data sets

ABSTRACT

This paper describes a normalized index for assessing the performance of PID control loops. The index was developed with a focus on ease of implementation by making use of features that are commonly found in modern control systems. Exponentially-weighted moving averages (EWMAs) or equivalent first-order infinite-impulse response (IIR) filters are found in many control system algorithm libraries and the index is developed using two EWMAs. Because the index is normalized, this paper shows how a single threshold can be derived from generic performance metrics thereby providing a solution to replace traditional setpoint alarming. Results are presented from evaluating the index using data gathered from several different loop types in a building heating, ventilating, and air-conditioning (HVAC) system.

1. Introduction

Modern control systems in buildings and in other types of applications are capable of monitoring and storing vast amounts of data. A common problem is that the amount of data can be overwhelming for operators and it can be difficult to transform big-data sets into useful information. Despite the trend toward advanced control such as model predictive control (MPC) and other model based strategies, PID (proportional, integral, and derivative action) feedback controllers are at the foundation of many control systems and their performance is thus critical to the overall performance of the system. It is therefore important to monitor and assess the performance of these controllers in order to ensure the integrity of the system as a whole.

Performance assessment of single loop feedback control is a mature field of research that can be traced back to the early stages of automatic control where the response of a control loop to disturbances such as a setpoint change or a load disturbance would be evaluated with measures such as the integrated absolute error (IAE) (Astrom, 1995). The difficulty with traditional metrics such as IAE is that they are not normalized and different loop types cannot be compared with each other. The seminal work of Harris solved this problem by proposing a performance metric based on measured variance normalized by the theoretical minimum variance for each loop (Harris, 1989). The Harris index allowed different loop types to be compared with each other and led to a surge in interest in control performance assessment. Several survey papers and books discuss developments in this area, e.g., Harris, Seppala, & Desborough,1999; Joe Qin, 1998; Jelali, 2006; Ordys, Uduehi, & Johnson, 2007; Gómez, Daniel, Moya, & Baeyens, 2010; Huang & Shah, 2012. Other examples of contributions to this area can be found in Horch and Isaksson (1999), Horch (2000) and Hägglund (1995).

In a large scale plant, there can be hundreds, if not thousands, of control loops with different characteristics. Normalized indexes enable loops of different types to be compared on the same scale thereby allowing quicker identification of poorly performing loops and allowing maintenance and supervision resources to be deployed more efficiently. There is a strong demand for PID performance metrics in buildings and HVAC (heating, ventilating, and air-conditioning) systems as well as in other low-cost application areas. However, adoption has been disappointing mainly because of the difficulty of implementation. Many of the methods popular in applications that are not as cost-sensitive require new algorithms to be developed and installed in existing control systems or data need to be recorded and stored for off-line processing. These barriers, although relatively minor, are enough to prevent wide-scale adoption of controller performance monitoring. The index described in this paper seeks to overcome these obstacles by making implementation a matter of utilizing existing functionality. Computation is very simple and the approach does not require storage of data. The core of the idea is to examine the symmetry of the process variable around setpoint. Other researchers have followed a similar approach to detect specific control loop problems such as stiction (Jelali & Huang, 2009).

In contrast to other normalized performance indexes, such as those based on minimum variance, the proposed index quantifies the setpoint error relative to its expected value of zero rather than the variance relative to an optimal benchmark. Although variance is a very

* Corresponding author.
E-mail address: timothy.i.salsbury@jci.com (T.I. Salsbury).

http://dx.doi.org/10.1016/j.conengprac.2016.12.002
Received 13 May 2016; Received in revised form 1 December 2016; Accepted 2 December 2016
0967-0661/

important factor for control performance in many industrial processes, it is less important than the setpoint error for applications such as HVAC. A persistent setpoint error usually indicates a more serious problem in a system than a higher level of variance and is thus useful in identifying faults that would often go undetected in low-cost applications where scant resources are devoted to monitoring. The proposed index thus approaches the control performance assessment problem by focusing on the first rather than second moment of a stochastic error signal. In this way, it is complementary with most of the other normalized performance indexes that have been proposed. It should be noted that, similar to other indexes such as those based on minimum variance, the proposed index is designed for regulatory control problems. Large setpoint changes could thus lead to false alarms in certain circumstances and should therefore be handled appropriately in a practical implementation.

The performance index described in this paper was first presented by the authors in Salsbury and Alcala (2015). The approach is extended here by developing a method for calculating a threshold for the index based on generic performance metrics. Additional results are also presented from tests carried out on three different loop types in an occupied building. The paper is structured as follows. First, the EWMA statistics that are required to calculate the index are described in Section 2 and recommendations are made for selecting the forgetting rate. The normalized index is then derived from two EWMAs in Section 3 and a method for selecting a generic threshold is presented in Section 3.1. Results from tests on different HVAC control loops in an occupied building are presented in Section 3.3 followed by conclusions in Section 4.

2. EWMA calculation

The index was developed with ease of implementation in mind by working backwards from the functionality already available in low-cost control systems. EWMAs or the equivalent infinite-impulse response (IIR) first-order filters are fairly ubiquitous in modern control systems and these form the basis of the proposed normalized index. EWMAs are also widely used in the area of quality control and statistical process control where they are often deployed to identify changes in the mean of a variable over time, e.g., Chao-Wen & Reynolds, 1999. EWMAs are exponentially-weighted moving averages with the most recent values getting the greatest weighting and the weighting decaying exponentially with time.

Because EWMAs are recursively calculated statistics, data storage requirements are small, thus placing little extra burden on the control system infrastructure. The proposed index is derived from two EWMA statistics with each one operating on the error signal (e) in the control loop, which is defined as the difference between the setpoint (r) and the controlled variable (y):

$$e_k = r_k - y_k \tag{1}$$

where k is a sample number beginning at a value of one. The two EWMA statistics are described below.

The first EWMA is calculated using the unmodified error signal samples as follows:

$$\mathrm{ewma}_{1,k} = \mathrm{ewma}_{1,k-1} + \frac{e_k - \mathrm{ewma}_{1,k-1}}{\min(k,\ W)} \tag{2}$$

The first EWMA is an exponentially-weighted average of the error signal, which under normal operation should have an expected value of zero. This EWMA is a statistical measure of location for the error signal. The second EWMA is calculated from the absolute value of the error signal:

$$\mathrm{ewma}_{2,k} = \mathrm{ewma}_{2,k-1} + \frac{|e_k| - \mathrm{ewma}_{2,k-1}}{\min(k,\ W)} \tag{3}$$

Because the second EWMA calculates an exponentially-weighted moving average of the *absolute* value of the error signal it represents a measure of statistical scale with the expected value implicitly assumed to be zero. The two EWMAs therefore provide measures of both scale and location of the error signal. Note that W is the effective number of samples used in the weighted averages. The forgetting time constant (τ_E) of the EWMAs is determined by the W value and also by the execution period (Δt) of the equation, i.e., $\tau_E = W\Delta t$. Thus, either W or Δt can be adjusted to achieve a different forgetting time constant. Note that the use of the minimum in the denominator of the update term causes the statistics to begin as straight averages until the number of samples reaches the window size when they become exponentially-weighted. Inspection of Eqs. (2) and (3) shows that the initial value of the EWMAs (i.e., at k=1) is not important because the previous values of the EWMA cancel from the equations when k=1. For example, for Eq. (2):

$$\mathrm{ewma}_{1,1} = \mathrm{ewma}_{1,0} + \frac{e_1 - \mathrm{ewma}_{1,0}}{1} \tag{4}$$

$$= e_1 \tag{5}$$

and, as stated earlier, the EWMA begins as a recursive mean while $k \le W$:

$$\mathrm{ewma}_{1,2} = e_1 + \frac{e_2 - e_1}{2} = \frac{1}{2}(e_1 + e_2) \tag{6}$$

$$\mathrm{ewma}_{1,n} = \frac{1}{n}(e_1 + e_2 + \cdots + e_n) \text{ for } n \le W \tag{7}$$

The second EWMA is a measure of dispersion and is related to the sample standard deviation of the error signal. For a normally distributed variable, its standard deviation can be calculated as a multiple of its mean absolute deviation (MAD); Appendix A.1 shows the calculation of this relationship.

To use the EWMAs to calculate the proposed index, the forgetting time constant should be set based on consideration of the open loop residence time (or time constant) of the plant under control. It is recommended that τ_E be set to five or more multiples of the plant residence time so that the EWMA statistics capture the behavior of the control loop over a long enough period to cover the range of potential aggressive and sluggish responses of the controller.

3. Normalized index

The normalized index is based on the assumption that the control loop is subject to disturbances that are drawn from a symmetrical probability distribution. This assumption is reasonable and is aligned with stochastic control theory where disturbances are often assumed to originate from a Gaussian distribution. Based on this assumption, the process variable can be expected to fluctuate above and below its setpoint. When a problem arises and the setpoint cannot be attained, this assumption will be violated. The proposed index can detect these kinds of problems even when the magnitude of the deviation from setpoint is small.

Traditional approaches to setpoint alarming usually involve defining thresholds above and below the setpoint. Small errors that are persistent will not be detected by these approaches even though they might indicate a more serious control defect. Another difficulty with traditional setpoint alarming is that momentary transgressions of alarm limits can generate unwanted alarms. This is often addressed by incorporating some type of timer so that an alarm is only generated when the error has persisted for a certain time. The proposed approach removes this additional complexity by naturally incorporating a time-based response into the index so that persistence of error as well as magnitude is handled holistically.

Analysis of the EWMA statistics described in the previous section reveals that ewma_1 should have an expected value of zero when the control is good and when deviations about setpoint are distributed