Event-triggered strategies for industrial control over wireless networks (invited paper)

Maben Rabi
ACCESS Linnaeus Centre
School of Electrical Engineering
Royal Institute of Technology
100 44 Stockholm, Sweden
maben.rabi@ee.kth.se

Karl H. Johansson
ACCESS Linnaeus Centre
School of Electrical Engineering
Royal Institute of Technology
100 44 Stockholm, Sweden
karl.henrik.johansson@ee.kth.se

ABSTRACT

New event-based sampling strategies can support the efficient use of radio resources in wireless control systems. Motivated by the recent introduction of wireless network nodes in process control industry, we consider the particular demands these closed-loop systems set on the wireless communication and the influence the communication has on the control performance. In the paper, it is pointed out that by letting sensor nodes transmit only when needed, it is possible to minimize the communication bandwidth utilization in these systems. We show how classical control strategies commonly based on periodic sampling, such as proportional-integral-derivative control and minimum variance control, can be cast in an event-based setting in which decentralized communication decisions are taken suitable for commonly used contention-based medium access control protocols. Event-triggered sampling for estimation is also reviewed. Simulated examples illustrate the results.

1. INTRODUCTION

There is a growing deployment of wireless networks in industrial control and automation. The lower installation cost and easier system reconfiguration for wireless devices can have a major influence on future control systems. Employing several control loops over a common wireless medium raises however new issues on how to allocate radio resources in an efficient way with guarantees on closed-loop system performance for the control applications. In some situations, a deterministic scheduling of the communication medium for control and estimation applications is required. In many cases, however, feedback control can provide good performance also with contention-based access schemes. In this paper, we show precisely that by indicating how event-triggered sensing and control provides a more scalable and efficient trade-off between control performance and communication cost. By making transmissions only when needed and taking the communication decisions locally at the sensor nodes, it is possible to minimize the communication bandwidth utilization.

Proportional-integral-derivative (PID) control is by far the dominating controller in process industry [5]. Design and implementation of PID controllers with periodically sampled sensor readings and control actuations is a well established area [4]. Stochastic optimal control is in its simplest formulation denoted minimum variance control, in which, the variance of the plant output is minimized by suitable zero-order hold control [2]. The traditional implementation of also minimum variance control is through periodic updates of the controller and actuator.

The main contribution of this paper is to propose extensions of these traditional control architectures to event-based implementations. By removing the constraint of periodic communication of sensor and control data, the wireless network resources can be used more efficiently. The proposed event-triggered control architecture is shown in Figure 3. An event detector decides on when to transmit plant information depending on the sensor measurements. The receiver generates a control command that is executed by the actuator. The main inspiration to this scheme comes from the work on event-triggered and time-triggered control in [3]. More recent contributions in the area includes [9, 11, 7, 10]. A deadband-based PID controller was proposed in [1].

The outline of the paper is as follows. Section 2 introduces event-triggered PID control. In particular, a suitable event
detector based on the desired proportional and integral action is introduced. Section 3 presents event-triggered sampling for minimum variance control, where it is shown that an event detector based on a time-varying threshold is optimal. Event-triggered sampling for estimation is reviewed in Section 4. Finally, the paper is concluded in Section 5. The results of Sections 3–4 are discussed in further detail in [10] and [11], respectively.

2. EVENT-TRIGGERED SAMPLING FOR PID CONTROL

In this section we present a new scheme for event-triggered PID control. We first recall the basics of PID control. Then we discuss the event detector and the control generator blocks in the event-triggered control architecture of Figure 3. The section is ended with an illustration of disturbance rejection for a first-order plant.

Consider the control system
\[ \dot{x}_1 = f(x_1, u_1, d_1), \]
where \( x_1 \) is the controlled state, \( u_1 \) the control, and \( d_1 \) an exogenous disturbance signal. The function \( f \) is supposed to be Lipschitz continuous. To simplify the discussion, suppose the system is scalar.

In a conventional PID controller, we form the error signal
\[ \xi_t = r_1 - x_1, \]
where \( r_1 \) is the reference signal to track and \( x_1 \) is the plant output. The error and the integral of the error
\[ dp_i = \int_0^t \xi_s ds, \]
form the basis of PID control action, together with the derivative of the error. We focus on PI control and recall that in continuous-time control systems, the PI controller has the basic structure
\[ u_t = -K_P \xi_t - K_I p_i. \]
In sampled data control, the sampling is typically periodic and the sampled data PI controller [4] is a discretization of the continuous-time PI controller. Within the sample intervals, a zero-order hold is usually employed, so the control signal is piecewise constant.

2.1 Event detector for PI control

The internal structure of the event detector block of the control architecture of Figure 3 is shown in Figure 2. The event detector consists of a P-part with a gain and a level crossing detector and an I-part with an integrator and another level crossing detector. The input to the event detector is the plant error and the output of the event detector is the encoded word to be communicated over the wireless medium.

The P- and the I-parts of the event detector generate the sampling instances that trigger the communication and thus the updates of the control generator. Let the sequence of sampling times be denoted
\[ \mathcal{T} = \{\tau_0, \tau_1, \ldots\}. \]

Let \( \tau_{\lambda} \) denote the last sample time, where \( \lambda \) is the count of samples till now. Some of these samplings are generated by the P-part and the others by the I-part.

P-sampling

The P-part monitors the instantaneous magnitude of the error signal \( \xi_t \). The idea is that, if the error grows in magnitude, we need to register this as an event to update the control action. We do this by looking for up crossing times of \( |\xi_t| \), where up crossings are level crossings away from zero. Down crossings of \( |\xi_t| \) mean that plant is behaving well.

Assume a threshold of \( \delta > 0 \) for the level crossing of the P-part. Define the P-sampling instances as the up-crossing times
\[ \theta_{i,\delta} = \inf \{t : |\xi_t| < \delta, t > \tau_i\}. \]
This sequence of crossing times capture two types of events. We trigger a sample if the unsigned error signal has grown by an amount of \( \delta \). If the error signal diminishes in magnitude since the last sample time \( \tau_i \), as desired, but has shot over the origin by at least \( \delta \), we also trigger a new sample.

It is not enough to use P control. It is easy to see that P control can result in a steady state error. To prevent such a situation, the sensor needs to check if the error signal is indeed decaying to zero. I-sampling describe next, addresses this concern.

I-sampling

Assume a fixed threshold \( \eta \) for changes in the integral error signal. For \( i = 1, 2, \ldots \), define the I-sampling instances as
\[ \pi_{i,\eta} = \inf \{t : |p_{i+1} - p_i| \geq \eta, t > \tau_i\}. \]
Note that we do not mind if the integral error is stuck at a particular value or if it is always between two crossing levels, because if the integral error converges, the error signal itself is converging to zero.

The sampling time succeeding \( \tau_i \) is the earlier of the P-sampling and the I-sampling times:
\[ \tau_{\lambda+1} = \min \{\theta_{i,\delta}, \pi_{i,\eta}\}. \]

If the system has reached the desired reference point and if it stays there, no further samples are triggered until either the reference changes or the system is hit by a fresh disturbance. If the thresholds \( \delta, \eta \) are increased, the communication cost is lowered to the cost of lower performance.

At sampling times, the control waveform is updated in the control generator block of Figure 3. We will now suggest a form for these updates.

![Figure 2: Event detector for PI control.](image-url)
2.2 Control generator for PI control

The control waveform we propose is a superposition of two simple waveforms. The two base waveforms are a series of pulses triggered by samples of the error signal, and a piecewise constant waveform whose values depend on samples of the integral error signal. Recall that the sampling times are determined by level crossings of either the error signal or of its integral. At these times, the parameters for updating the control waveform are transmitted to the controller.

The control waveform is given by:

\[ u_t = \begin{cases} -K_P \xi - K_I \rho, & \text{for } t \in [\tau_\lambda, \tau_\lambda + \epsilon(K_P, K_I)], \\ -K_I \rho, & \text{for } t \in (\tau_\lambda + \epsilon(K_P, K_I), \tau_{\lambda+1}), \end{cases} \]

where \( \epsilon(K_P, K_I) \) is the duration for which the proportional part of the update in control is applied. The integral component is on the other hand applied till the next update is received. This control waveform is inspired by the traditional PI controller, but has updates only at event-triggered time instances. It is an improvement over the purely proportional event-triggered controller.

2.3 Disturbance rejection

Next we illustrate how the event-triggered PID controller acts by applying it to the scalar linear plant:

\[ dx_t = -\frac{1}{T} x_t + Ku_t + d_t, \]

where \( K > 0 \) is the gain and \( T > 0 \) the time constant of the plant. The disturbance \( d_t \) is an exogenous non-deterministic signal. We assume that it is an unknown piecewise constant signal, which is a common and realistic assumption in practice. The task of a good controller is to make sure that the disturbance is attenuated at the plant output.

Traditional PI control relies on integral action to counter disturbances. Our event-triggered PI controller is able to capture this merit of periodic PI control. Figures 3–5 show simulations of the closed-loop system when the reference \( r_t \) is zero. The parameters of the plant are \( T = 1 \) and \( K = 1 \). The system is at steady state when, at time zero, a unit step disturbance is applied to the plant input. To isolate the behavior of the integral action, we set \( K_P \) equal zero.

Figure 3 shows that with a threshold for the integral error \( \eta = 1 \) and unit control increment, the event-triggered controller regulates the error to zero with a single sampling event. If instead \( \eta = 0.5 \) and the control increment is 0.1, more sampling events are needed as shown in Figure 4. Finally, Figure 5 indicates that if not careful tuning of the integral part is done, sustained oscillations can happen. Systematic tuning rules for the event-triggered PID control are currently being developed. In particular, we look into the problem of how to transfer a PID control design for a periodic implementation to a event-triggered implementation based on the control architecture of Figure 3.

3. EVENT-TRIGGERED SAMPLING FOR MINIMUM VARIANCE CONTROL

In the previous section we discussed how PID control can be implemented on the event-triggered architecture of Figure. Next we present how the solution to the stochastic
control problem of minimum variance control can be cast into the same framework. In particular, we suppose that the output of the control generator is piecewise constant and we search for an event detector that minimize the output of the control generator is piecewise constant into the same framework. In particular, we suppose that control problem of minimum variance control can be cast.

3.1 Joint optimal control and time stopping problem

On the time interval \([0,T]\), consider a scalar controlled diffusion process \(x_t\) governed by the stochastic differential equation

\[
dx_t = \mu(x_t, t, u_t)dt + \sigma(x_t, t, u_t)dB_t,
\]

where \(B_t\) is a standard Brownian motion process, the functions \(f\) and \(g\) are Lipschitz continuous, and the control process \(u_t\) satisfies the usual causality property of being measurable w.r.t. \(x_t\), thus ensuring existence and uniqueness of \(x_t\). We assume that the sensor measures the state continuously and with negligible measurement noise. The control process is given by the piecewise constant signal

\[
u_t = U_N \cdot 1_{\{\tau_N \leq t < T\}} + \sum_{i=0}^{N-1} U_i \cdot 1_{\{\tau_i \leq t < \tau_{i+1}\}},
\]

where \(\tau_i, i = 0, \ldots, N\), are the so-called stopping times with respect to \(x_t\). These time instances correspond to the events generating the communication over the wireless network. Figure 6 illustrates a case with \(N = 1\) switch instance for the piecewise constant control. We seek a feedback control policy and sampling events that minimize the objective function

\[
J_{\text{General}} = \mathbb{E} \left[ \int_0^T c(x_s) ds \right],
\]

where the function \(c(\cdot)\) is non-negative. This is a joint optimal control and time stopping problem, since the decision variables are both the controls \(\{U_i\}_{i=0}^N\) and the stopping times \(\{\tau_i\}_{i=0}^N\). Note that the decision on the controls are to be taken in the control generator while the stopping times are to be decided on in the event detector. We will see that unlike in the case of periodic sampled-data control, there is no separation between optimal estimation and control for this problem.

3.2 Special case: controlled Brownian motion

We consider the special case of controlled Brownian motion with a quadratic objective function and one switch time, which is the simplest instance of the general problem discussed above. It is also a case when the solution can be explicitly computed in closed form. Consider the particular case of (3) given by

\[
dx_t = dB_t + u_t dt.
\]

Assume the process starts at the origin, \(x_0 = 0\). The piecewise constant control is given by

\[
u_t = \begin{cases} U_0 = U_0(x_0, T) & \text{if } 0 \leq t < \tau, \\ U_1 = U_1(x_\tau, T - \tau) & \text{if } \tau \leq t \leq T, \end{cases}
\]

see Figure 6. The objective function is given by

\[
J = \mathbb{E} \left[ \int_0^T x_s^2 ds \right].
\]

The decision variables are thus \(U_0, U_1\), and \(\tau\). The computation of their optimal values \(U_0^*, U_1^*, \) and \(\tau^*\) is illustrative and thus summarized next.

Since no decision is taken in the interval \((\tau, T)\), it is easy to derive the optimal choice of the terminal control level \(U_1^*\).

It is given by the linear feedback law

\[
U_1^* (x_\tau, T - \tau) = \frac{3x_\tau}{2(T - \tau)},
\]

where hence \(U_1^*\) depends on the state at the switching instance and how close the switching instance is to the end time \(T\). The optimal choice \(U_0^*\) on the other hand, is closely tied with the the choice of \(\tau\). In fact, for the best control performance, \(U_0^*\) and \(\tau\) have to be picked jointly. By using the expression for \(U_1^*\), we have that minimizing the objective function (5) is achieved by minimizing

\[
J_0 = \mathbb{E} \left[ \int_0^T x_s^2 ds \right] + \mathbb{E} \left[ \frac{1}{4} x_\tau^2 (T - \tau) + \frac{(T - \tau)^2}{2} \right].
\]

with respect to \(U_0\) and \(\tau\). It is intuitive that \(U_0^* = 0\) because the Brownian motion is a martingale with zero mean, so it would be costly to steer the mean of the process away from the origin using nonzero \(U_0\). This gives

\[
J_0 = \frac{T^2}{2} + \frac{U_0^2 T^4}{3} - \mathbb{E} \left[ \left( \sqrt{\frac{x_\tau}{3}} + \frac{U_0 (T - \tau)}{\sqrt{3}} \right)^2 (T - \tau) \right]
\]

\[
= \frac{T^2}{2} - \frac{3}{4} \mathbb{E} \left[ x_\tau^2 (T - \tau) \right].
\]

It remains to optimize \(J_0\) with respect to \(\tau\). The only crucial restriction on \(\tau\) is on the causal dependence on the state trajectory. It follows from the expression of \(J_0\), that we seek an admissible stopping time \(\tau\) that maximizes the expected reward function

\[
\mathbb{E} \left[ x_\tau^2 (T - \tau) \right]
\]
This can be done using Itô calculations [8]. Suppose that we can find a smooth function \( g(x, t) \) for which
\[
\begin{align*}
g(x, t) & \geq x^2 (T - t) \quad \text{(6)} \\
0 & = \frac{1}{2} g_{xx} + g_t, \quad \text{(7)}
\end{align*}
\]
where the inequality is tight in the sense that for every \( T - t \), there exists at least one \( x \) for which the inequality becomes an equality. Then, by the Itô change of variable formula [8], we have that for every stopping time \( \tau \) with bounded expectation and \( 0 \leq t \leq \tau \leq T \),
\[
E \left[ x^2 (T - \tau) \mid F_\tau \right] \leq E \left[ g(x, \tau) \right],
\]
which gives the Snell envelope
\[
\tau^* = \inf \left\{ t \mid x^2 (T - t) \right\}.
\]
Thus at any time \( t \) on the interval \([0, T]\), we have \( g(x, t) \) as an upper bound for the expected reward for stopping at a later time. Moreover, as we will see shortly, this bound is achievable so that it is also the maximum achievable expected reward. The process \( S_t := g(x, t) \)
is called the Snell envelope for the optimal stopping problem. Because of the assumption on \( g \), the inequality (6) connecting the instantaneous reward function and the function \( g \) is tight. Hence, there exists a stopping time with bounded expectation which will achieve the upper bound for the maximum expected reward. This time is precisely the earliest instant when the reward collected by stopping equals the Snell envelope:
\[
\tau^* = \inf \left\{ t \leq \nu \leq T \mid g(x, \nu) = x^2 (T - \nu) \right\}.
\]

We will now furnish an explicit solution to the sufficiency conditions (6)–(7) for the controlled Brownian motion. Consider as the candidate solution the fourth Hermite polynomial of the ratio \( x_t / \sqrt{3} \) multiplied by the scaling factor \( \sqrt{3} / 2 \). Then,
\[
g(x_t, t) = \sqrt{3} / 2 \left( \frac{x_t}{\sqrt{3}} \right)^4 + \left( \frac{x_t}{\sqrt{3}} \right)^2 + \frac{1}{2},
\]
which gives the Snell envelope
\[
S_t = A \left( \frac{x_t^4}{6} + x_t^2 (T - t) + \frac{(T - t)^2}{2} \right).
\]

We require the inequality
\[
S_t = g(x_t, t) \geq x_t^2 (T - t),
\]
to be tight. Indeed the choice \( A = \sqrt{3} / 1 + \sqrt{3} \) makes the quantity \( h := g - x^2 (T - t) \) a perfect square, thus providing the desired tightness:
\[
h(x_t, t) = \frac{1}{2 \left( 1 + \sqrt{3} \right)} \left( \frac{x_t^4}{3} - 2 \frac{x_t^2}{\sqrt{3}} (T - t) + (T - t)^2 \right).
\]

Thus the optimal stopping rule is the symmetric quadratic envelope
\[
\tau^* = \inf \left\{ t \mid x^2 (T - t) \right\}.
\]

Figure 7 shows the optimal envelope together with an example of a state trajectory and corresponding optimal control. The envelope is to be implemented in the event detector. The optimal deterministic switching is to switch at \( T/2 \). The expected control performance cost incurred by the optimal event-triggered switching scheme is \( T^2/8 \), which is nearly a third of the the cost of using deterministic switching.

4. EVENT-TRIGGERED SAMPLING FOR ESTIMATION

In this section we consider an event-triggered estimation problem. This work was previously reported in [11].

We consider here an estimation problem for the Brownian motion process \( x_t \). The state is sampled at instants \( \{ \tau_i \}_{i \geq 0} \) which are stopping times w.r.t. the \( x \)-process. We will now describe the MMSE estimate and its variance. Consider a time \( t \) in the (semi-open) interval \([\tau_i, \tau_{i+1})\). We have:
\[
\hat{x}_t = E \left[ x_t \mid \tau_i \leq t < \tau_{i+1}, \left\{ (\tau_j, x_{\tau_j}) \right\}_{0 \leq j \leq i} \right],
\]
\[
= E \left[ x_t \mid \tau_i \leq t < \tau_{i+1}, \tau_i, x_{\tau_i} \right].
\]

Similarly, its variance \( p_t \) can be written as:
\[
\hat{x}_t = E \left[ (x_t - \hat{x}_t)^2 \right] \leq \tau_{i+1} - \tau_i, x_{\tau_i} \right].
\]

It can be proved [9] that the optimal MMSE estimate reduces to
\[
\hat{x}_t = x_{\tau_i} \quad \forall \ t \in [\tau_i, \tau_{i+1}).
\]

The task then is to choose the sampling times so that the
following estimation distortion is minimized:

$$J(T, N) = \mathbb{E} \left[ \int_0^T (x_s - \hat{x}_s)^2 ds \right].$$

We first state the solution to the single sampling problem which is the case when the sample budget equals one. In the previous section on minimum variance control, we obtained a solution for the optimal single sampling time problem which minimized the ensuing control cost when the initial state is zero. The single sampling problem for estimation has the sample optimal stopping rule as the control problem with zero initial state [10]. The minimum estimation distortion with one allowed sample can be computed to be $3T^2/8$.

Now, we will inductively obtain a parametric expression for the minimal expected distortion given exactly $k$ samples with $k > 1$. Consider the stopping cost:

$$J(\tau) = \mathbb{E} \left[ \int_{0}^{\min(\tau, T)} x_s^2 ds + \frac{\alpha}{2} (T - \tau)^2 \right] \tag{9}$$

where $\alpha \geq 0$ is a given constant. We can rewrite this as:

$$\frac{1}{2} \left( T^2 - \mathbb{E} \left[ 2x_{\min(\tau, T)}^2 (T - \tau) + (1 - \alpha)[(T - \tau)^2] \right] \right).$$

Like in the single sample case, let us pay attention to the part of the above expression which depends on $\tau$ and define the following optimal stopping problem:

$$\min_{\tau} \mathbb{E} \left[ 2x_{\tau}^2 (T - \tau) + (1 - \alpha)(T - \tau)^2 \right].$$

Consider the candidate maximum expected reward function:

$$g(x, t) = A \left\{ (T - t)^2 + 2x^2 (T - t) + \frac{x^4}{3} \right\},$$

where $A$ is a constant chosen such that $g(x, t) - 2x^2(T - t) - (1 - \alpha)(T - t)^2$ becomes a perfect square. The only possible value for $A$ then is:

$$(5 + \alpha) - \sqrt{(5 + \alpha)^2 - 24}.$$ 

Then the optimal stopping time is given by:

$$\tau^* = \inf \left\{ t : g(x_t, t) \leq 2x_t^2(T - t) + (1 - \alpha)(T - t)^2 \right\},$$

and the corresponding optimal distortion $\mathcal{J}$ becomes

$$\mathcal{J} = (1 - A)\frac{T^2}{2}.$$ 

Now, we obtain the explicit stopping rules and the corresponding minimal distortions for different values of the sample budget $N$ by defining recursively $\kappa_N, \gamma_N$:

$$\kappa_N = 1 - \frac{(5 + \kappa_{N-1}) - \sqrt{(5 + \kappa_{N-1})^2 - 24}}{4},$$

$$\gamma_N = \frac{3\kappa_{N-1} - \kappa_N}{1 - \kappa_N}.$$ 

The $(k + 1)^{th}$ sampling time is chosen as:

$$\tau_{k+1} = \inf_{t \geq \tau_k} \left\{ t : (x_t - x_{\tau_k})^2 \geq \gamma_{N-k+1}T - t \right\}.$$ 

Figure 8: Estimation error signal and the optimal envelopes when exactly three samples of a standard Brownian motion process are to be generated. The envelope for any sampling time $\tau$ is parabolic: $x = r_{\tau - \sqrt{T - \tau}}$, where, $r_1 = 1.3161$, $r_2 = 0.8819$, $r_3 = 0.6953$.

5. CONCLUSIONS

Event-based sampling strategies that support the efficient use of radio resources in wireless networks in feedback control systems were discussed. It was shown that some classical controllers, like PID control and minimum variance control, which are typically implemented using periodically sampled data, can be implemented using an aperiodic sampling scheme. The new control structure was based on an event detector at the sensor side that decides when to transmit information to the controller node. Examples were given when an optimal event detector can be analytically derived. Optimal estimation for event-based sampling was also briefly presented.

6. ACKNOWLEDGMENTS

This research was partially funded by the Swedish Research Council, the Swedish Governmental Agency for Innovation Systems, Swedish Foundation for Strategic Research, and the European Commission through SOCRADES and FeedNetBack.

7. REFERENCES


