Recovery of delay time from time series based on the nearest neighbor method

M.D. Prokhorov a,*, V.I. Ponomarenko a,b, V.S. Khorev b

a Saratov Branch of Kotel’nikov Institute of Radio Engineering and Electronics of Russian Academy of Sciences, Zelyonyaya Street, 38, Saratov 410019, Russia
b Department of Nano- and Biomedical Technologies, Saratov State University, Astrakhanskaya Street, 83, Saratov 410012, Russia

A R T I C L E   I N F O

Article history:
Received 12 August 2013
Accepted 29 September 2013
Available online 4 October 2013
Communicated by C.R. Doering

Keywords:
Parameter estimation
Nonlinear time-delayed feedback systems
Time series analysis
Nearest neighbors

A B S T R A C T

We propose a method for the recovery of delay time from time series of time-delay systems. The method is based on the nearest neighbor analysis. The method allows one to reconstruct delays in various classes of time-delay systems including systems of high order, systems with several coexisting delays, and nonscalar time-delay systems. It can be applied to time series heavily corrupted by additive and dynamical noise.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Self-sustained oscillators with delay-induced dynamics are highly widespread in nature. Their abundance results from such fundamental features as the finite velocity of signal propagation that is especially displayed in spatially extended systems [1] and time-delayed feedback inherent in many physical [2,3], chemical [4], climatic [5], and biological [6–8] systems and processes. Studying time-delay systems it is important to know the delay times whose values in many respects define the system dynamics and features. Knowledge of delay times is of considerable significance in model construction and prediction of system behavior in time and under parameter variation. That is why the problem of delay time reconstruction from experimental time series attracts a lot of attention.

To solve this problem a variety of methods has been proposed, which allows one to recover the delay times of time-delayed feedback systems from their chaotic time series. Many of these methods are based on the projection of the infinite-dimensional phase space of time-delay systems onto low-dimensional subspaces [9–14]. They use different criteria of quality for the system reconstruction, for example, the minimal forecast error of the constructed model [9–11], minimal value of information entropy [12], or various measures of complexity of the projected time series [13,14]. The methods of delay time recovery are known based on employment of regression analysis [15,16], statistical analysis of time intervals between extrema in the time series [17], information-theory approaches [18,19], multiple shooting approach [20], optimization algorithm [21], and adaptive synchronization [22,23]. A separate group of methods for delay time estimation is based on the analysis of the time-delay system response to external perturbations [24–27]. These methods can be applied to systems performing not only chaotic, but also periodic oscillations.

In this Letter we propose a novel method for recovering delay time from time series. It is based on the nearest neighbor method. The method of nearest neighbors is widely used in different scientific disciplines for nonlinear time series analysis [28–31]. Its main areas of application are classification of objects and forecast of time series. In the object classification problem the basic idea of the nearest neighbor method is that the object is assigned to the class of its nearest neighbor or to the class most common amongst its k nearest neighbors. In application to the forecast of a time series the method idea is to use for prediction of a future state of a system its states in the past, which are most similar to the current state. We propose using the nearest neighbor method for the first time for estimating the delay time of a delayed feedback system from time series.

The Letter is organized as follows. In Section 2 we present the idea of the method and apply it to recover first-order time-delay systems with a single delay in chaotic and periodic regimes. In Sections 3 and 4 the method is applied for the reconstruction of delays in scalar time-delay systems of second order and with several coexisting delays, respectively. We show that the proposed method can be used for determining an a priori unknown order
of the model equation and the number of delays. In Section 5 the method is applied for the recovery of delay time in nonscalar time-delay system. In Section 6 we summarize our results.

2. Recovery of delay time in first-order time-delay systems with a single delay

Let us explain the method idea with one of the most popular first-order delay-differential equation with a single delay:

$$\varepsilon \dot{x}(t) = -x(t) + f(x(t-\tau)),$$  \hspace{1cm} (1)

where $\tau$ is the delay time, the parameter $\varepsilon$ characterizes the inertial properties of the system, and $f$ is a nonlinear function. Note that the Mackey–Glass equation [6] and the Ikeda equation [1], which become standard equations in the study of time-delay systems, can be reduced to Eq. (1).

Analyzing time series, we always deal with variables measured at discrete instants of time. Therefore, it is convenient to pass from differential Eq. (1) to the difference equation

$$\frac{x(t + \Delta t) - x(t)}{\Delta t} = -x(t) + f(x(t-\tau)), \quad (2)$$

where $\Delta t$ is the sampling time. Eq. (2) can be rewritten as

$$x(t + \Delta t) = a_1 x(t) + a_2 f(x(t-\tau)), \quad (3)$$

where $a_1 = 1 - \Delta t/\varepsilon$ and $a_2 = \Delta t/\varepsilon$. Let us write Eq. (3) in the form of the discrete-time map

$$x_{n+1} = a_1 x_n + a_2 f(x_{n-d}), \quad (4)$$

where $n = t/\Delta t$ is the discrete time and $d = \tau/\Delta t$ is the discrete delay time.

Assume that we have a time series $\{x_n\}_{n=1}^N$ from the system (1), where $N$ is the number of points. Let us define vector $\vec{X}_i = (x_i, x_{i-d})$ and find vector $\vec{X}_j = (x_j, x_{j-d})$ with $j \neq i$, which is a nearest neighbor of $\vec{X}_i$. The nearest neighbor for a given vector can be chosen according to some metrics [30]. The most widely used metrics is the Euclidean metrics

$$L(\vec{X}_i, \vec{X}_j) = \sqrt{ (x_i - x_j)^2 + (x_{i-d} - x_{j-d})^2}. \quad (5)$$

The vector $\vec{X}_j$ will be the nearest neighbor of $\vec{X}_i$, if the distance $L(\vec{X}_i, \vec{X}_j)$ is minimal. Generally, it is a common practice to find not one, but $k$ nearest neighbors for a given vector.

The basic idea of the proposed method is that the nearest neighbor vectors containing the system (4) dynamical variable at the instants of time $n$ and $n-d$, where $n \in \{d+1, N-1\}$, will lead to the close states of the system at the instants of time $n+1$, because the system (4) evolution is defined by its current state and the state at the delayed instant of time. Since the delay time is a priori unknown, we vary the trial delay times $m$ within some interval and for $k$ nearest neighbors of each vector $\vec{X}_m = (x_m, x_{m-d})$ constructed from the time series estimate the variance $\sigma_n^2$ of the system states at the corresponding instants of time $n+1$.

In the case of false choice of $m(m \neq d)$, the variance of these states may be great, because the system states at the instants of time $n+1$ do not depend on the system states at the instants of time $n-m$. True delay time $d$ can be estimated as the value at which the minimum of the following dependence:

$$D(m) = \frac{1}{N-m-2} \sum_{n=m+1}^{N-1} \sigma_n^2 \quad (6)$$

is observed.

We apply the method to time series of the Mackey–Glass equation

$$\dot{x}(t) = -bx(t) + \frac{ax(t-\tau)}{1 + x^a(t-\tau)}, \quad (7)$$

which can be converted to Eq. (1) by division by $b$. The parameters of Eq. (7) are chosen to be $a = 0.2$, $b = 0.1$, $c = 10$, and $\tau = 300$ to produce a dynamics on a chaotic attractor. The sampling time is $\Delta t = 1$ and the number of points is $N = 10000$. Part of the time series is shown in Fig. 1(a).

Fig. 1(b) depicts the dependence of $D$ on the trial delay time $m$ for different numbers $k$ of nearest neighbors for vector $\vec{X}_m = (x_m, x_{m-d})$. The value of $m$ is varied from 1 to 500 with a step of 1. All the dependences $D(m)$ exhibit a well-pronounced absolute minimum at $m = 300$, which provides an accurate recovery of the discrete delay time $d = \tau/\Delta t = 300$.

If the time series points are sampled with a high frequency, a situation is possible in which the vectors $\vec{X}_j = (x_j, x_{j-d})$ with $j = i \pm p$ ($p = 1, 2, \ldots, P$) that are close in time to vector $\vec{X}_i = (x_i, x_{i-d})$ will be detected as its nearest neighbors. To avoid this undesirable situation in the search for the nearest neighbors of vector $\vec{X}_i = (x_i, x_{i-d})$, one should exclude from consideration $l = 2P$ vectors $\vec{X}_j = (x_j, x_{j-d})$ close to $\vec{X}_i$ in time.

The dependences $D(m)$ are plotted in Fig. 1(c) for $k = 10$ and different numbers $l$ of close in time vectors, which are not taken into account in searching for nearest neighbors. All the plots...
exhibit a sharp absolute minimum at \( m = d = 300 \), as well as the plots in Fig. 1(b).

It should be noted that instead of searching for a fixed number \( k \) of nearest neighbors for vector \( \vec{X}_i = (x_i, x_{i-d}) \), one can assign all vectors \( \vec{X}_j = (x_j, x_{j-d}) \) to its nearest neighbors, if \( L(\vec{X}_i, \vec{X}_j) < \delta \), where \( \delta \) is a small quantity. The plots of \( D(m) \) constructed in this way of finding nearest neighbor vectors are similar to the plots presented in Fig. 1(b). The appropriate choice of the parameters \( k \) and \( \delta \) enables one to achieve almost complete coincidence of the results of searching for nearest neighbors in both ways. In addition, we have found that the choice of the metrics for searching nearest neighbors has almost no effect on the form of the dependences \( D(m) \).

To test the method efficiency in the presence of noise we apply it to the data produced by adding a zero-mean Gaussian white noise to the time series of Eq. (7). The obtained results are presented in Fig. 2(a) for different levels of additive noise at \( k = 10 \) and \( l = 10 \). The location of the minimum of \( D(m) \) allows us to recover the delay time accurately even for noise level of about 30% (the signal-to-noise ratio is about 10 dB). Such level of noise greatly exceeds the noise level that is allowed for applying most of other methods of delay time reconstruction.

The proposed method is even more robust with respect to the dynamical noise. In Fig. 2(b) the dependences \( D(m) \) are shown at \( k = 10 \) and \( l = 10 \) for the case, where a zero-mean Gaussian white noise is added to the right-hand side of Eq. (7). In all the plots constructed in Fig. 2(b) for different levels of noise the minimum of \( D(m) \) is observed at \( m = 300 \).

Let us consider the case where the system (7) performs periodic oscillations (\( a = 0.2 \), \( b = 0.1 \), \( c = 10 \), and \( \tau = 10 \)). Part of the time series of these oscillations is shown in Fig. 3(a). In the construction of the dependences \( D(m) \) we consider vectors \( \vec{X}_i \) as nearest neighbors of vector \( \vec{X}_j \), if \( L(\vec{X}_i, \vec{X}_j) < 0.02 \). In the absence of noise there is no pronounced minimum in the plot of \( D(m) \) (Fig. 3(b)). However, the presence of dynamical noise turns out to be useful for recovering the delay time. A clear minimum appears in the plot of \( D(m) \) at \( m = d = 10 \) in the case of a 10% noise (the signal-to-noise ratio is 20 dB) (Fig. 3(b)). Certainly, the presence of additive noise has no positive effect on determining the delay time.

In contrast to most of other methods for the reconstruction of delay time, the proposed method can be applied to estimating the delay time not only in the systems (1) characterized by a linear dependence from the current state and nonlinear dependence from the delayed state. It can be also applied to time series gained from more general type of time-delay systems with nonlinear function \( F \) depending on both variables \( x(t) \) and \( x(t-\tau) \):

\[
\dot{x}(t) = r x(t) \left( 1 - \frac{x(t-\tau)}{q} \right). \tag{8}
\]

The reasoning presented above for the system (1) holds for the system (8), since the proposed method takes into account the dependence of evolution of the system on its current state and the state at the delayed instant of time, while the form of this dependence is of no importance.

Let us apply the method to time series of the Hutchinson system [32]:

\[
\dot{x}(t) = r x(t) \left( 1 - \frac{x(t-\tau)}{q} \right). \tag{9}
\]

At \( r = 1.7 \), \( q = 1 \), and \( \tau = 10 \) the system (9) performs periodic oscillations. We added a zero-mean Gaussian white noise to the right-hand side of Eq. (9) and constructed the dependence \( D(m) \). Fig. 4 presents the plot of \( D(m) \) for \( k = 10 \), \( l = 10 \), and a 5% dynamical noise (the signal-to-noise ratio is about 26 dB). For the sampling time \( \Delta t = 1 \) the minimum of \( D(m) \) is observed at \( m = 10 \), which coincides with the delay time \( d = \tau / \Delta t = 10 \).

3. Recovery of delay time in second-order time-delay systems

The proposed method can be easily extended to high-order time-delay systems. In particular, it can be modified for the systems described by the second-order delay-differential equations

\[
\varepsilon_2 \ddot{x}(t) + \varepsilon_1 \dot{x}(t) = F(x(t), x(t-\tau)), \tag{10}
\]

where \( \varepsilon_1 \) and \( \varepsilon_2 \) are the parameters characterizing the inertial properties of the system. As an example we consider the following system:

\[
\varepsilon_2 \ddot{x}(t) + \varepsilon_1 \dot{x}(t) = -x(t) + f(x(t-\tau)). \tag{11}
\]

Using the described above formalism, one can pass from differential Eq. (11) to the discrete-time map

\[ x_{n+2} = b_1 x_{n+1} + b_2 x_n + b_3 f(x_{n-d}). \tag{12} \]
where \(b_1 = 2 - (\varepsilon_1 \Delta t)/\varepsilon_2\), \(b_2 = -1 + (\varepsilon_1 \Delta t - (\Delta t)^2)/\varepsilon_2\), and \(b_3 = (\Delta t)^2/\varepsilon_2\).

For each vector \(\vec{x}_n = (x_{n+1}, x_n, x_{n-m})\) constructed from Eq. (11) time series we find \(k\) nearest neighbor vectors and estimate for them the variance \(\sigma^2_d\) of the system states at the corresponding instants of time \(n + 2\). Then we calculate the dependence

\[
D(m) = \frac{1}{N - m - 3} \sum_{n=m+1}^{N-2} \sigma^2_d
\]

under variation of the trial delay time \(m\). The location of the minimum of (13) will give us an estimation of the discrete delay time \(d = \tau/\Delta t\).

The proposed methods can be used for determining an a priori unknown order of a delayed feedback system from its time series. To define the order of the time-delay system one has to recover initially its delay time under the assumption that the system is described by the first-order equation (1). Then, one has to recover the delay time under the assumption that the system model equation is the second-order equation (11) and construct the dependences (6) and (13) in the same plot. The dependence \(D(m)\) constructed under the true choice of the model equation order will lie below the dependence \(D(m)\) constructed under the false choice of the order of the model equation.

For example, let us have a time series from the second-order time-delay system (11) with quadratic nonlinearity and \(\Delta t = 1\) that is slightly larger than \(\tau = 10\). It has a minimum at \(m = 999\) and lies mainly higher than the black curve indicating that the model equation has the form of Eq. (11) is shown in Fig. 6 in grey color. It has a minimum at \(m = 999\) and lies mainly higher than the black curve indicating that the model equation of the system has the first order.

4. Recovery of delay times in time-delay systems with two delays

The proposed method can be also extended to systems with multiple delays. Let us consider a time-delay system with two different delay times \(\tau_1\) and \(\tau_2\):

\[
\dot{x}(t) = -\varepsilon(t) + f_1(x(t - \tau_1)) + f_2(x(t - \tau_2)).
\]

Using the described above approach, we pass from differential Eq. (14) to the discrete-time map

\[
x_{n+1} = a_1x_n + a_2 f_1(x_{n-d_1}) + a_2 f_2(x_{n-d_2}),
\]

where \(a_1 = 1 - \Delta t/\varepsilon\), \(a_2 = \Delta t/\varepsilon\), \(d_1 = \tau_1/\Delta t\), and \(d_2 = \tau_2/\Delta t\).

From Eq. (15) it follows that the nearest delay vectors containing the dynamical variable at the instants of time \(n, n - d_1, \) and \(n - d_2\), where \(n \in [d_2 + 1, N - 1]\) \((d_2 > d_1)\), will lie to the close states of the system at the instants of time \(n + 1\). Since the delay times \(d_1\) and \(d_2\) are unknown, we vary the trial delay times \(m_1\) and \(m_2\) within some interval and for \(k\) nearest neighbors of each vector \(\vec{x}_n = (x_n, x_{n-m_1}, x_{n-m_2})\) constructed from the time series estimate the variance \(\sigma^2_d\) of the system states at the corresponding instants of time \(n + 1\).

In the case of false choice of \(m_1\) and/or \(m_2\) \((m_1 \neq d_1\) and/or \(m_2 \neq d_2\)), the variance of these states may be great. The location of the minimum of the dependence

\[
D(m_1, m_2) = \frac{1}{N - m_2 - 2} \sum_{n=m_2+1}^{N-1} \sigma^2_d
\]

can be used as an estimation of the delay times \(d_1\) and \(d_2\).

We demonstrate the method efficiency with a generalized Mackey–Glass equation obtained by introducing a further delay,

\[
\dot{x}(t) = -bx(t) + \frac{1}{2} \frac{ax(t - \tau_1)}{1 + x^e(t - \tau_1)} + \frac{1}{2} \frac{ax(t - \tau_2)}{1 + x^e(t - \tau_2)}.
\]

Division of Eq. (17) by \(b\) reduces it to Eq. (14) with \(\varepsilon = 1/b\). For \(a = 0.2, b = 0.1, c = 10, \tau_1 = 70,\) and \(\tau_2 = 300\) Eq. (17) exhibits chaotic oscillations. The dependence \(D(m_1, m_2)\) representing...
a two-dimensional surface is shown in Fig. 7(a) for $N = 10000$ and $\Delta t = 1$. In the construction of this figure the vectors $\vec{X}_i$ and $\vec{X}_j$ were considered as nearest neighbors, if the distance

$$L(\vec{X}_i, \vec{X}_j) = \sqrt{(x_i - x_j)^2 + (x_{i-m_1} - x_{j-m_1})^2 + (x_{i-m_2} - x_{j-m_2})^2}$$

was less than 0.02. The dependence $D(m_1, m_2)$ has a global minimum at $m_1 = d_1 = 70$ and $m_2 = d_2 = 300$ providing an accurate recovery of both delay times.

For comparison Fig. 7(b) presents the dependence $D(m)$ constructed by applying the method proposed in Section 2 for the system (1) with a single delay to the time series of Eq. (17). The plot of this dependence described by Eq. (6) shows deep minima at $m = 68$ and $m = 298$. Hence, the delay time estimation appears to be less accurate without taking into account the form of model Eq. (14). One more distinctive minimum of $D(m)$ is observed in Fig. 7(b) close to $m = d_1 + d_2$.

Note that applying the method of two delays reconstruction to time series of Eq. (1) with a single delay, we observed the minimum of $D(m_1, m_2)$ at $m_1 = m_2 = d$.

The method described in this section can be applied to the recovery of delay times not only in the systems (14), but also in more general class of systems with two delays governed by equation

$$\dot{x}(t) = F(x(t), x(t - \tau_1), x(t - \tau_2)).$$

5. Recovery of delay time in nonscalar time-delay systems

The method for the recovery of delay time from time series based on the nearest neighbor analysis can be extended to nonscalar time-delay systems

$$\dot{x}(t) = F(x(t), x(t - \tau), y(t - \tau)), \quad \dot{y}(t) = F(y(t), y(t - \tau)).$$

In this case, using the time series of both variables $x(t)$ and $y(t)$ one has to search for the nearest neighbors for vectors $\vec{X}_n = (x_n, x_{n-m}, y_{n-m})$ or $\vec{Y}_n = (y_n, y_{n-m}, x_{n-m})$ under variation of trial delays $m$ and determine the delay time by the location of the minimum of the dependence (6).

Let us apply the method to time series of a system of two coupled nonlinear delayed equations

$$\dot{x}(t) = r x(t) - \mu [x^2(t - \tau) + cy^2(t - \tau)] x(t),$$

$$\dot{y}(t) = r y(t) - \mu [x^2(t - \tau) + cy^2(t - \tau)] y(t)$$

introduced in [33]. We choose the parameters to be $r = 4$, $\mu = 4$, $c = 0.5$, and $\tau = 0.35$. As it was shown in [33], at these parameter values the system (21) shows periodic oscillations. Part of the time series of $x(t)$ is presented in Fig. 8(a).

To construct the plot of $D(m)$ we use the time series of $x(t)$ and $y(t)$ with $N = 10000$ and $\Delta t = 0.01$. In the absence of noise the dependence $D(m)$ has no pronounced minimum. However, the dependence $D(m)$ shows a minimum, if independent dynamical noises with sufficiently large intensity are added to the right-hand side of both equations in the system (21). In Fig. 8(b) the dependence $D(m)$ is constructed for $k = 10$, $l = 10$, and 40% dynamical noises (the signal-to-noise ratio is about 8 dB). It has a minimum at $m = 35$, which coincides with the delay time $d = \tau / \Delta t = 35$.

6. Conclusion

We have proposed the method for the reconstruction of delay time in time-delay systems from their time series. The method is based on the nearest neighbor analysis. It allows one to recover the delay times in scalar time-delay systems of different order and with multiple delays and nonscalar time-delay systems. The method can be applied to time-delay systems with arbitrary form of nonlinear function, including the function depending on both the delayed and non-delayed variables. Moreover, the method can be used for determining an a priori unknown order of a time-delay system from its time series. The parameters of the method can be chosen within a wide range.

The proposed method remains efficient under very high levels of dynamical and additive noise. It is shown that the method can be successfully applied to the recovery of delay time in time-delay systems performing chaotic oscillations and time-delay systems performing periodic oscillations in the presence of dynamical noise.

Acknowledgements

This work is supported by the Russian Foundation for Basic Research, Grant No. 13-02-00227.

References