Computers and Chemical Engineering 36 (2012) 294–300

Contents lists available at SciVerse ScienceDirect

Computers and Chemical Engineering

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

Online soft sensor for hybrid systems with mixed continuous and discrete measurements

Evgenia Suzdaleva [a,*], Ivan Nagy [b]

[a] Department of Adaptive Systems, Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 18208 Prague, Czech Republic

[b] Faculty of Transportation Sciences, Czech Technical University, Na Florenci 25, 11000 Prague, Czech Republic

A R T I C L E   I N F O

Article history:
Received 17 December 2010
Received in revised form 28 June 2011
Accepted 7 September 2011
Available online 21 September 2011

Keywords:
Online state prediction
Hybrid filter
State-space model
Mixed data

A B S T R A C T

Online state prediction and fault detection are typical tasks in the chemical industry. In practice it often happens that some variables, important and critical for quality control, cannot be measured online due to such restrictions as cost and reliability. An uncertainty existing in real systems allows to use a probabilistic approach to online state estimation. Such an approach is proposed in this paper. Different types of information appearing in an online diagnostic system are processed via combination of algorithms subject to probability distributions. This combination of algorithms is presented as a decomposed version of Bayesian filtering. In this paper, the proposed solution is specialized for a system with mixed both continuous and discrete-valued measurements and unobserved variables.

1. Introduction

Online state prediction and fault detection are typical tasks in the chemical industry. In practice it often happens that some variables, important and critical for quality control, cannot be measured online due to such restrictions as cost and reliability. It means that the needed variables must be estimated using a model concerned with an available set of measurements (for example, temperature in non-isothermal reactors, gaseous flow rates, dissolved oxygen concentration in bioreactors, etc.).

Many various approaches (model-driven as well as data-driven soft sensors) are developed for state prediction in the field of industrial processing plants (e.g., Lin, Recke, Renaudat, Knudsen, & Jörgensen, 2005; Park & Han, 2000). Their main task is to provide estimates of unmeasured variables based on the knowledge of the process dynamics and on the available online observations. A detailed overview of the most popular soft sensor techniques is given in Kadlec, Gabrys, and Strandt (2009). However, despite the high number of papers, there are still open or weakly supported issues in this area.

Many tasks in the process industry (in chemical plants as well) still need manual control and decision-making of an operator. An online state predictor supports the operators and allows them to make faster and more objective decisions. Another problem is the rather dynamic nature of the processing plants. It means that a soft sensor should react fast to sudden process input changes, which is a difficult task, usually involving high costs.

Fault detection and diagnosis methods are nicely described in Venkatasubramanian, Rengaswamy, Kavuri, and Yin (2003). According to Venkatasubramanian et al. (2003), design and implementation of nonlinear model-based soft sensors are still limitedly supported and limitedly reliable for chemical processes. Most quantitative methods are based on input–output models, and when they are restricted to linear applications, their advantages over statistical technique such as Principal Component Analysis (PCA) are minimal. Besides nonlinearity, modeled quantities can demonstrate both a continuous and a discrete character (for example, the system is normal or faulty, it is one or another fault, etc.).

Research published in Kadlec et al. (2009) and Venkatasubramanian et al. (2003) show that there is no single universal method to handle all the requirements for a diagnostic system. Single-method based systems are seriously limited in application, which again increases the necessity of manual supervising. Combination of methods in the form of a hybrid system is seen as a flexible adaptive way. Exploitation of different types of knowledge in a hybrid system will contribute to fast and more effective decision-making, and finally, integration of diagnostic methods with other process operations will lead to a more comprehensive intelligent supervisory control system.

An uncertainty existing in real systems allows to use a probabilistic approach to online state estimation. Such an approach is proposed in this paper. Different types of information appearing

* Corresponding author.
E-mail addresses: suzdalev@utia.cas.cz (E. Suzdaleva), nagy@utia.cas.cz (I. Nagy).

0098-1354/$ – see front matter
doi:10.1016/j.compchemeng.2011.09.004

in an online diagnostic system are processed via combination of algorithms subject to probability distributions. This combination of algorithms is presented as a decomposed version of Bayesian filtering (Kárný et al., 2005). Universal in terms of probability (density) functions, here it focuses on a system with mixed both continuous and discrete-valued measurements and unobserved variables. It means both linear and nonlinear information is modeled. The proposed method uses analytical solutions as far as possible. It means that numerical procedures are applied only in those parts which cannot be computed analytically.

Layout of the paper is as follows. Section 2 describes problem formulation, basic facts about models used and Bayesian filtering. Section 3 is devoted to probabilistic approach to online hybrid filtering. Section 4 demonstrates application of general solution to normal and discrete multinomial models. Section 5 provides illustrative examples with real discrete data sample and normal simulations. Comparison with one of the counterparts is presented. Remarks in Section 6 close the paper.

2. Preliminaries

2.1. Problem formulation

The system considered is concerned with the following observations: the system output y_t and the control input u_t measured at discrete time moments $t=\{1,\ldots,T\}\equiv t^*$. In general, all the variables are column vectors such as $y_t=[y_{1;t},\ldots,y_{Y;t}]'$, $u_t=[u_{1;t},\ldots,u_{U;t}]'$ including entries that can be both continuous and discrete-valued.

The system state $x_t=[x_{1;t},\ldots,x_{X;t}]'$ is not directly observed and has to be estimated in an online (*recursive*) mode.

2.2. State-space model

The system is described by the state-space model in the form of the following conditional probability (density) functions (p(d)fs)

$$\textit{observation model}: \quad f(y_t|x_t,u_t), \tag{1}$$

$$\textit{state evolution model}: \quad f(x_{t+1}|x_t,u_t), \tag{2}$$

for simplicity denoted as pdfs within this paper.

The unobserved state can be estimated with the help of Bayesian filtering (Kárný et al., 2005).

2.3. Bayesian filtering

Bayesian filtering includes the following coupled formulas.

Data updating

$$f(x_t|D(t)) = \frac{f(y_t|x_t,u_t)f(x_t|D(t-1))}{\int f(y_t|x_t,u_t)f(x_t|D(t-1))\,dx_t} \propto f(y_t|x_t,u_t)f(x_t|D(t-1)) \tag{3}$$

incorporates information contained in observations $D(t)=(d_1,\ldots,d_t)$, where $d_t\equiv(y_t,u_t)$. Relation (3) also comprises the natural conditions of control (Peterka, 1981), according to those

$$f(x_t|u_t,D(t-1))=f(x_t|D(t-1)).$$

Time updating

$$f(x_{t+1}|D(t)) = \int f(x_{t+1}|x_t,u_t)f(x_t|D(t))\,dx_t \tag{4}$$

fulfills state prediction.

The prior pdf $f(x_1|D(0))$ which expresses the subjective prior knowledge of the system's initial state starts the recursions.

2.4. Chain rule

An operation intensively used throughout the paper is *Chain rule*

$$f(a,b|c)=f(a|b,c)f(b|c) \tag{5}$$

which decomposes the joint pdf $f(a,b|c)$ into a product of conditional pdfs for any random variables a, b and c.

3. Recursive hybrid filtering

Let's consider a system with observations $y_t=[y_t^c,y_t^d]'$, $u_t=[u_t^c,u_t^d]'$, and with an unobserved state to be estimated $x_t=[x_t^c,x_t^d]'$, where superscript c denotes a continuous type of variable, while superscript d belongs to a discrete variable. Sign $'$ denotes transposition. All the variables in general are multivariate.

To perform estimation for such a hybrid system, firstly, Bayesian filtering (3) and (4) is proposed to be done in one integration step, i.e.,

$$f(x_{t+1}|D(t)) \propto \int f(x_{t+1}|x_t,u_t)\left\{\underbrace{f(y_t|x_t,u_t)f(x_t|D(t-1))}_{\propto f(x_t|D(t))}\right\}dx_t, \tag{6}$$

which is obtained by a trivial substitution of the state estimate updated by measurements (3) into the time updating (4).

The basic idea of the proposed hybrid filter consists in application of the chain rule (5) to models (1) and (2) and relation (6). All the joint pdfs must be decomposed to products of conditional pdfs of the individual vector entries. This decomposition allows to model and estimate variables which bring different-type information, i.e., they can be either continuous or discrete in their values.

3.1. Decomposition of models

Using the chain rule (5), one decomposes the observation model (1)

$$\begin{aligned} f(y_t|x_t,u_t) &= f(y_t^c,y_t^d|x_t^c,x_t^d,u_t^c,u_t^d) \\ &= f(y_t^c|y_t^d,x_t^c,x_t^d,u_t^c,u_t^d)f(y_t^d|x_t^c,x_t^d,u_t^c,u_t^d) \\ &= f(y_t^c|y_t^d,x_t^c,u_t^c)f(y_t^d|x_t^d,u_t^d) \end{aligned} \tag{7}$$

realistically assuming that the past discrete state and discrete input can be omitted from the condition for the continuous output y_t^c, and continuous entries – from the condition for discrete y_t^d. The obtained decomposition represents a product of pdfs.

Similarly the state evolution model (2) is decomposed as

$$\begin{aligned} f(x_{t+1}|x_t,u_t) &= f(x_{t+1}^c,x_{t+1}^d|x_t^c,x_t^d,u_t^c,u_t^d) \\ &= f(x_{t+1}^c|x_{t+1}^d,x_t^c,x_t^d,u_t^c,u_t^d)f(x_{t+1}^d|x_t^c,x_t^d,u_t^c,u_t^d) \\ &= f(x_{t+1}^c|x_{t+1}^d,x_t^c,u_t^c)f(x_{t+1}^d|x_t^d,u_t^d). \end{aligned} \tag{8}$$

3.2. Prior pdfs

A form of the prior pdf should be specified to be used in (6). The joint prior pdf is also decomposed as

$$f(x_t|D(t-1))=f(x_t^c,x_t^d|D(t-1))=f(x_t^c|x_t^d,D(t-1))f(x_t^d|D(t-1)). \tag{9}$$

Preserving this product form even for posterior distribution is a condition for recursive updating. This form is spoiled during estimation and must be approximately restored. This is the main theoretical result of the presented paper.

3.3. Hybrid state estimation

Substituting (7)–(9) in (6), one obtains the following relation

$$f(x^c_{t+1}|x^d_{t+1}, D(t))f(x^d_{t+1}|D(t))$$
$$\propto \int_{x^{c*}} \sum_{x^{d*}} \underbrace{f(x^c_{t+1}|x^d_{t+1}, x^c_t, u^c_t)f(x^d_{t+1}|x^d_t, u^d_t)}_{f(x_{t+1}|x_t,u_t)}$$
$$\times \underbrace{f(y^c_t|y^d_t, x^c_t, u^c_t)f(y^d_t|x^d_t, u^d_t)}_{f(y_t|x_t,u_t)} \underbrace{f(x^c_t|x^d_t, D(t-1))f(x^d_t|D(t-1))}_{prior\ pdf} dx^c_t$$
$$= \sum_{x^{d*}} f(x^d_{t+1}|x^d_t, u^d_t)f(y^d_t|x^d_t, u^d_t)f(x^d_t|D(t-1))$$
$$\times \int_{x^{c*}} f(x^c_{t+1}|x^d_{t+1}, x^c_t, u^c_t)f(y^c_t|y^d_t, x^c_t, u^c_t)f(x^c_t|x^d_t, D(t-1))\, dx^c_t. \quad (10)$$

Applying Bayesian filtering (3) and (4), and respectively (6), to the integral in (10) one obtains the updated state estimate $f(x^c_{t+1}|x^d_{t+1}, D(t))$ inside a sum over set of the discrete state values x^{d^*}. Similar application of (3) and (4) and (6) to discrete models in (10) with replacing the integration by regular summation gives the updated discrete state distribution, i.e., $f(x^d_{t+1}|D(t))$. In this way, the originally desired relation in (10) loses the prescribed form of the posterior pdf and becomes a sum of distributions, i.e.,

$$f(x^c_{t+1}|x^d_{t+1}, D(t))f(x^d_{t+1}|D(t)) \propto \sum_{x^{d*}} f(x^d_{t+1}|D(t))f(x^c_{t+1}|x^d_{t+1}, D(t)). \quad (11)$$

It is necessary to restore the original form (9) in order to use it for the next step of estimation. An approximation based on Kerridge inaccuracy is an explicit solution, which restores the original form of the pdf via computation of a specific weighted combination of the pdfs involved in (11). Kerridge inaccuracy (Kerridge, 1961) is a part of Kullback–Leibler divergence (Kullback & Leibler, 1951) adopted as a theoretically justified proximity measure. This divergence is known to be an optimal tool within the Bayesian approach (Kárný et al., 2005). For any random variable a, Kerridge inaccuracy is used to measure the proximity of pdfs $f(a)$ and $\hat{f}(a)$

$$K_a(f(a)||\hat{f}(a)) = \int_{a*} f(a) \ln \frac{1}{\hat{f}(a)}\, da \quad (12)$$

and its minimization allows to find the approximated pdf $\hat{f}(a)$. According to this approximation (Kárný et al., 2005), the sum in (11) is replaced by the product

$$\hat{f}(x^c_{t+1}|x^d_{t+1}, D(t))f(x^d_{t+1}|D(t)), \quad (13)$$

which is used as the prior pdf for the next step of recursive estimation.

The proposed general probabilistic approach assumes its universality in the sense of exploited distributions. Let's demonstrate its application to linear normal and discrete multinomial models.

4. Solution for normal and multinomial distributions

4.1. Models

The observation model (7) represents a product of distributions $f(y^c_t|y^d_t, x^c_t, u^c_t)f(y^d_t|x^d_t, u^d_t)$, where

$$f(y^c_t|y^d_t, x^c_t, u^c_t) = (2\pi)^{-Y/2}|R_v|^{-1/2}$$
$$\times \exp\left\{-\frac{1}{2}[y^c_t - Cx^c_t - Hu^c_t]' R_v^{-1}[y^c_t - Cx^c_t - Hu^c_t]\right\} \quad (14)$$

is the multivariate normal distribution, where C and H are known parameters, u^c_t is a known input and v_t is white Gaussian noise with zero mean and known covariance R_v. The multinomial distribution

$$f(y^d_t|x^d_t, u^d_t) \quad (15)$$

is provided by the output transition table and a known probability $\alpha_{q|l,n}$ with multi-index $q|l, n$. This multi-index denotes realizations $q \in \{1, \ldots, Q\}$ of random variable y^d_t at time instant t according to a set of its possible values $\{1, \ldots, Q\}$, where Q is a finite number. Realization q in the multi-index $q|l, n$ is conditioned by realizations $l \in \{1, \ldots, L\}$ of state x^d_t and $n \in \{1, \ldots, N\}$ of input u^d_t from their sets of possible values with finite numbers L and N. Notation $\alpha_{q|l,n}$ reflects probability of transition to the output $y^d_t = q$ conditioned by $x^d_t = l$ and $u^d_t = n$. It holds

$$\sum_{q=1}^{Q} \alpha_{q|l,n} = 1 \quad \text{and} \quad \alpha_{q|l,n} \geq 0 \quad \forall q, l, n.$$

Distribution (14) exists for each value q of discrete output y^d_t.

Similarly, the state evolution model (8) is a product $f(x^c_{t+1}|x^d_{t+1}, x^c_t, u^c_t)f(x^d_{t+1}|x^d_t, u^d_t)$, where

$$f(x^c_{t+1}|x^d_{t+1}, x^c_t, u^c_t) = (2\pi)^{-X/2}|R_w|^{-1/2}$$
$$\times \exp\left\{-\frac{1}{2}[x^c_{t+1} - Ax^c_t - Bu^c_t]' R_w^{-1}[x^c_{t+1} - Ax^c_t - Bu^c_t]\right\} \quad (16)$$

is the multivariate normal distribution, where A and B are known parameters and w_t is white Gaussian noise with zero mean and known covariance R_w; and

$$f(x^d_{t+1}|x^d_t, u^d_t) \quad (17)$$

is the multinomial distribution presented by the state transition table containing known probabilities $\beta_{l|m,n}$ with a multi-index $l|m, n$. Here the multi-index is evolved in a similar way as for the observation model but the condition $m \in \{1, \ldots, L\}$, which relates to the value of the state x^d_t at time instant t, while l here belongs to x^d_{t+1}. It holds

$$\sum_{l=1}^{L} \beta_{l|m,n} = 1 \quad \text{and} \quad \beta_{l|m,n} \geq 0 \quad \forall l, m, n.$$

Distribution (16) exists for each possible value of discrete state x^d_t.

4.2. Choice of prior distributions

The product of prior pdfs (9) is specialized in the following way

$$f(x^c_t|x^d_t, D(t-1))f(x^d_t|D(t-1)) \sim \mathcal{N}(\mu_t, P_t)p_{x^d_t}, \quad (18)$$

where the first factor $\mathcal{N}(\mu_t, P_t)$ denotes the prior normal distribution with the initial mean value μ_t and the initial covariance matrix P_t that has to be estimated for time instant $t+1$. The second factor $p_{x^d_t}$ is the prior multinomial distribution of discrete state x^d_t. Here it has the form of a vector containing the initial probabilities $p_l\, \forall l \in \{1,$

Table 1
Estimation error and correct point estimates.

	EE	CPE
HF	0.1117	216
IMM	0.1565	160

. . ., $L\}$ at time instant t, and it has to be estimated for time $t+1$. It holds

$$\sum_{l=1}^{L} p_l = 1 \quad \text{and} \quad p_l \geq 0 \quad \forall l.$$

4.3. State estimation for normal and multinomial distributions

In general, application of (3) and (4) to linear normal state-space model provides Kalman filter (Grewal & Andrews, 2001). Solution of the integral in (10) for normal distributions without calculation of the normalizing constant computationally coincides with the Kalman filter (Peterka, 1981), i.e.,

$$KG = P_t C'(C P_t C' + R_v)^{-1}, \tag{19}$$

$$\overline{\mu}_t = \mu_t + KG(y_t^c - C\mu_t - Hu_t^c), \tag{20}$$

$$\overline{P}_t = P_t - P_t C'(R_v + C P_t C')C P_t, \tag{21}$$

$$\mu_{t+1} = A\overline{\mu}_t + Bu_t^c, \tag{22}$$

$$P_{t+1} = A\overline{P}_t A' + R_w, \tag{23}$$

run for each value of x_t^d. This predicts the normal state providing its mean μ_{t+1} with covariance matrix P_{t+1}.

A part of (10) outside the integral corresponds to discrete multinomial distributions. The explicit solution of recursive state estimation for them takes the following form. With observations $y_t^d = q \in \{1, \ldots, Q\}$ and $u_t^d = n \in \{1, \ldots, N\}$ available at time instant t the predicted probability p_l for time instant $t+1$ is computed $\forall l \in \{1, \ldots, L\}$ in the following way

$$p_l = \beta_{l|1n}\alpha_{q|1n}p_1 + \beta_{l|2n}\alpha_{q|2n}p_2 + \cdots + \beta_{l|Ln}\alpha_{q|Ln}p_L \tag{24}$$

and then normalized, i.e.,

$$p_l = \frac{p_l}{\sum_{l=1}^{L} p_l},$$

resulting in the multinomial distribution

$$f(x_{t+1}^d | D(t)) = p_{x_{t+1}^d}, \tag{25}$$

which preserves the original form.

4.3.1. Approximation

However, the estimate of the normal state does not keep its original form, and relation (11) in this case is the mixture distribution $\sum_{l=1}^{L} p_l \mathcal{N}(\mu_{l,t+1}, P_{l,t+1})$, where $\mu_{l,t+1}$ and $P_{l,t+1}$ denote results of the Kalman filter (19)–(23) obtained for each value l. Restoring the original normal form needs to use the approximation based on Kerridge inaccuracy (Kerridge, 1961). According to Kárný et al. (2005), in the case of normal pdfs the Kerridge inaccuracy (12) is minimized with the following mean and covariance matrix of the approximated distribution

$$\hat{\mu}_{t+1} = \sum_{l=1}^{L} p_l \mu_{l,t+1}, \tag{26}$$

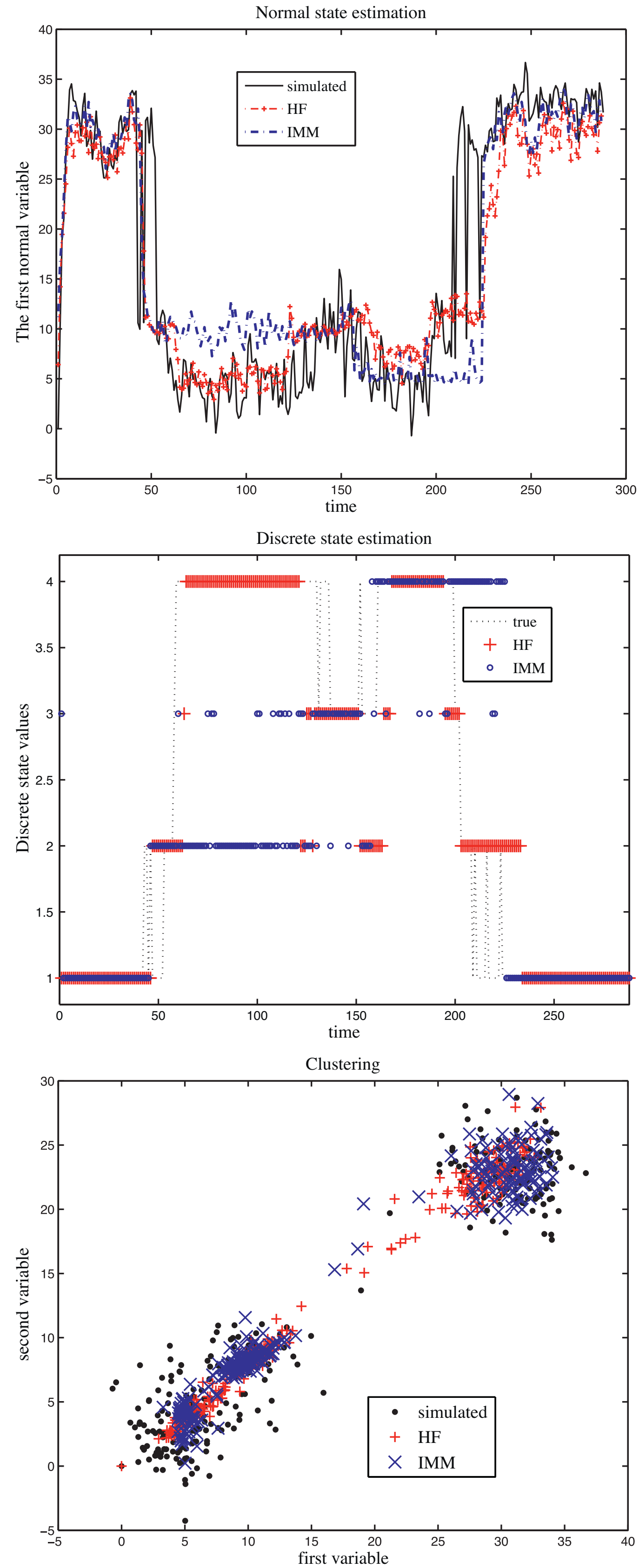

Fig. 1. Hybrid and IMM filters with higher valued diagonal covariance matrices. The top plot shows that the HF estimates are closer to simulated values than the IMM ones. At the middle plot the difference between results is more significant: the HF results cover most discrete values, while IMM does not catch value 4 from 70 to 120 time periods and value 2 around 215 time periods. The bottom plot shows the working modes found as clusters.

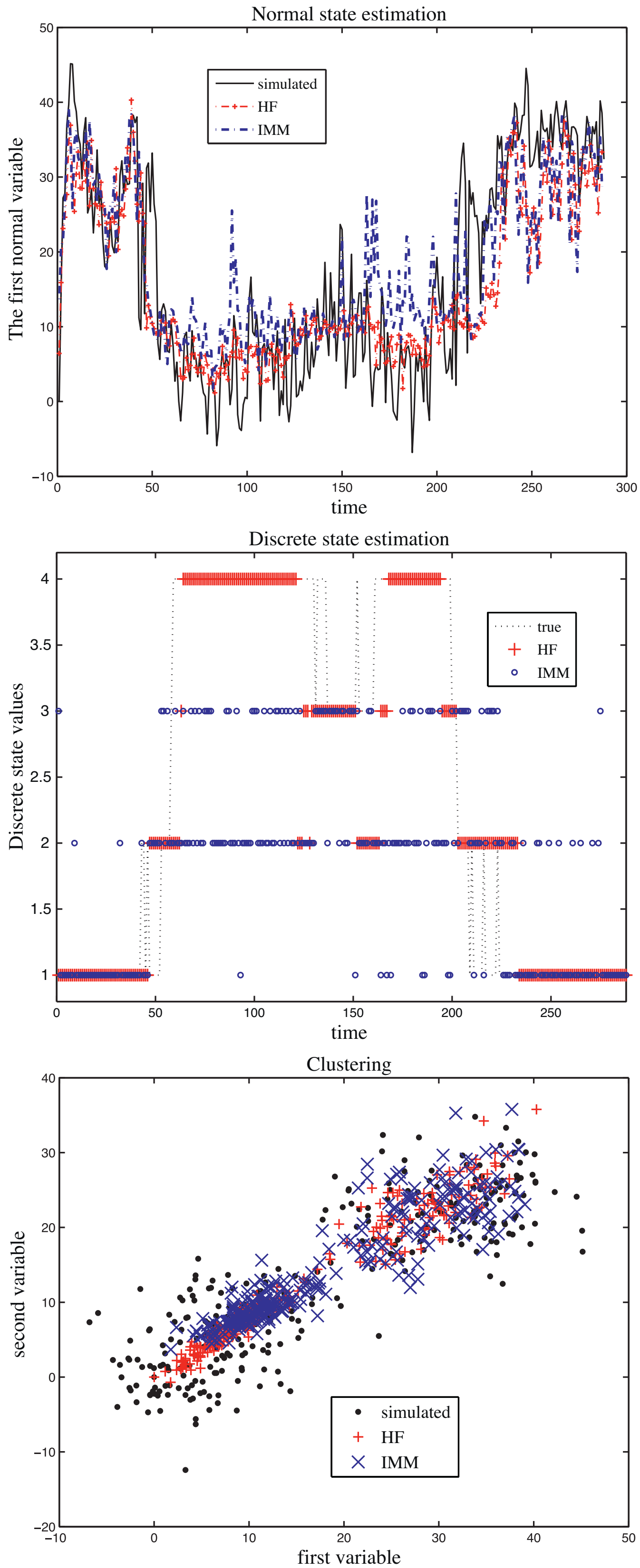

Fig. 2. Hybrid and IMM filters with full higher valued covariance matrices. Here, the noticeable advantage of HF at the middle plot can be seen: value 4 is not covered by IMM at all.

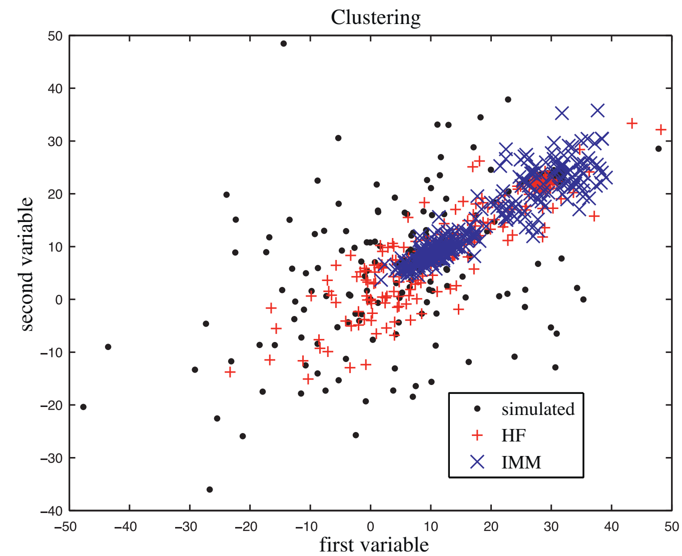

Fig. 3. Hybrid and IMM filters with a mixed character of noises. The difference among simulated data and estimates of both the methods is shown in the form of clustered working modes of the system.

$$\hat{P}_{t+1} = \sum_{l=1}^{L} p_l P_{l,t+1} + \sum_{l=1}^{L} p_l (\hat{\mu}_{t+1} - \mu_{l,t+1})^2. \tag{27}$$

The approximation (26)-(27) is then used as the prior normal distribution for the next step of the recursion.

5. Results

To test the proposed hybrid filter (HF), a real discrete data sample and simulations of the normal state-space model have been taken. The discrete data sample contains the state and the output both with 4 possible values and the control input with 2 possible values. The normally distributed two-dimensional variables have been simulated for 4 values of the discrete state that can be interpreted as the system working modes. This hybrid system state has been estimated via the proposed method and, to compare, with the help of the interactive multiple model (IMM) algorithm (Bar-Shalom, Kirubarajan, & Li, 2002). The IMM filter is a well-known approach, which performs Kalman filter (Grewal & Andrews, 2001) for each model and then computes a weighted combination of updated state estimates produced by all the filters. The IMM filter is close to that proposed in this paper. Difference is that the presented method takes the state-space model in a general form for both the normal and discrete states along with mixed observations and control inputs. Comparison of these filters provided the following results.

5.1. Experiment with diagonal covariance matrices and big noises

The first experiments were performed with rather noisy data, i.e., with higher-valued diagonal covariance matrices R_v and R_w. They were used both for simulation and estimation. The prior distributions were chosen in the same form for all experiments for both the algorithms. The estimation error (EE) was calculated in the following way

$$\mathrm{EE} = \frac{1}{T} \sum_{t} (x_t^c - \hat{\mu}_{t+1})'(x_t^c - \hat{\mu}_{t+1}),$$

where $T = 288$ is the duration of the estimation and x_t^c is the simulated state. For discrete state estimation, a number of correctly

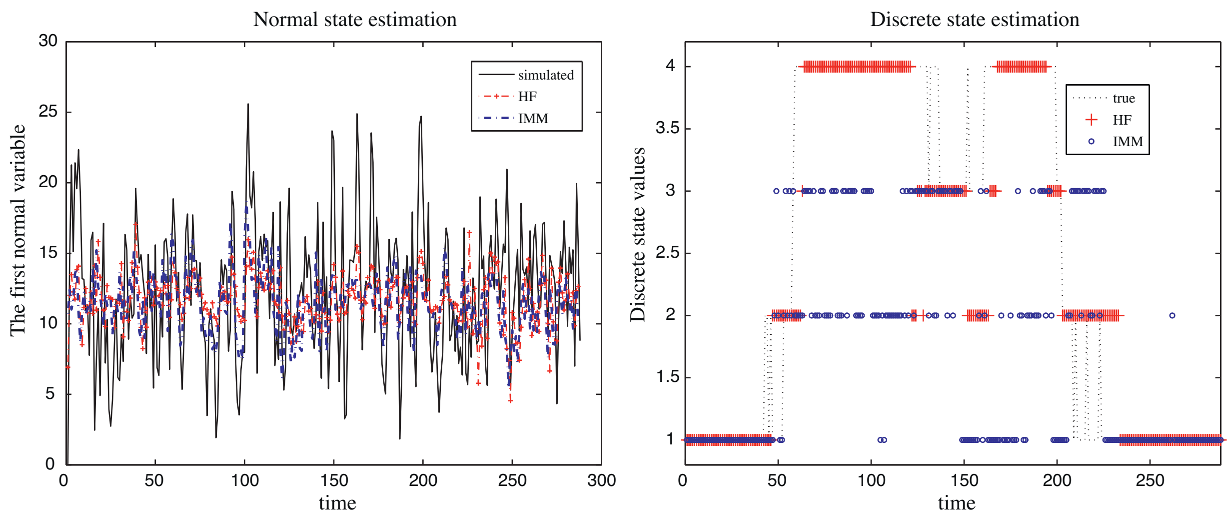

Fig. 4. Hybrid and IMM filters with difficult for distinguishing normal models. One can see in the left plot that the normal state estimates for both the filters differ insignificantly. The right plot shows a better stability of the hybrid filter: 215 correctly point-estimated states for the hybrid filter and 97 for the IMM one.

Table 2
Estimation error and correct point estimates.

	EE	CPE
HF	0.0980	216
IMM	∞	125

point-estimated states (CPE) from the total 288-data sample was evaluated for both the filters. It is assumed that for better quality of estimation, EE should have a minimal value, and CPE on the contrary – a maximal (from 288) value. EE and CPE can be seen in Table 1. They both demonstrate a better stability of the hybrid filter, although the difference between EE is rather small. Fig. 1 demonstrates comparison of the HF and IMM results for one of the normal state variables (top), discrete state (middle) and clustering (bottom). The estimation of the second normal state is of a similar quality. It can be seen that the advantage of the hybrid filter is more significant for the discrete state estimation.

5.2. Experiment with diagonal covariance matrices and small noises

The same data but with low-valued diagonal covariance matrices R_v and R_w, i.e., small noise systems, were used for this experiment. The IMM filter behaved unstably and the estimation failed. This was caused by the likelihood computation in the output prediction of the update step. The occurred problem of covariance matrix singularity seems to be a sensitive point of the IMM filter because of the great difference between output values during switching among working modes. Here low-noise data contributed to greater output value differences than in the previous experiments. The hybrid filter was stable with improved results in comparison with Section 5.1. The results can be found in Table 2.

5.3. Experiment with full covariance matrices and high noises

This experiment gave results similar to the first one with the advantage of the hybrid filter, see Table 3 and Fig. 2.

5.4. Experiment with mixed low and high noises

This experiment was performed with a mixed character of noises: small noises for some of the normal models and higher noises for others. Another random generator seed was used for

Table 3
Estimation error and correct point estimates.

	EE	CPE
HF	0.2012	216
IMM	0.2235	112

Table 4
Estimation error and correct point estimates.

	EE	CPE
HF	4.0918	215
IMM	3.5256	210

Table 5
Estimation error and correct point estimates.

	EE	CPE
HF	0.5965	216
IMM	0.8114	151

simulation. Here a sharp improvement of the IMM filter was observed that can be seen in Table 4. However, slight reduction of values of noise covariance matrices (keeping a mixed nature) again gave an improvement of HF, see Table 5. The working modes of the system as clusters found by both the filters are shown in Fig. 3.

This experiment motivated the comparison of filters from the point of view of "closeness" or "distinguishing" of normal models. In practice, the system working modes can be close to each other or even be partially overlapping. Such data is difficult to identify.

5.5. Experiment with very close and very different normal models

The experiment with very close and difficult for distinguishing normal models demonstrated a better stability of the hybrid filter: 215 correctly point-estimated states for the hybrid filter and 97 for the IMM one. The results are shown in Fig. 4. The normal state estimates for both the filters were obtained with a practically insignificant difference.

Results of the experiments with the sharply different models, see Table 6, and those with one sharply different normal model and others difficult for distinguishing, see Table 7, demonstrated a significantly better quality of the discrete state estimation for HF against IMM. Estimation quality for the normal states is also better

Table 6
Estimation error and correct point estimates.

	EE	CPE
HF	0.1355	216
IMM	0.1683	135

Table 7
Estimation error and correct point estimates.

	EE	CPE
HF	0.3036	217
IMM	0.3921	132

for the hybrid filter, although a difference in EE is not as significant as in CPE.

5.6. Discussion

To summarize the experimental part of the work, one can note that the hybrid filter was stable in all the experiments and had better results in that part which deals with estimation of the discrete state. In comparison to the IMM filter, the HF one is not sensitive to a greater difference in output values when a system is switching among working modes. Due to the explicit solution of discrete state estimation, computations in HF are much simpler. It contributed to the improved stability of HF, while the IMM estimation sometimes failed.

Moreover, it was observed during the experiments that the IMM filter gives better results when normal components are simulated with sharply different parameters, and therefore can be easier identified. The closer and more similar normal models, the worse the IMM results. The hybrid filter is more successful in that sense and does not possess this property due to the used approximation. It demonstrates a much smoother difference between its good and bad results.

6. Conclusion

The paper proposes the decomposed version of Bayesian filtering specialized for hybrid dynamic systems with normal and discrete multinomial states and observations. A similar problem is considered, for instance, in Elliot and Sworder (1996), where discrete state is treated via hidden Markov models (HMM) and a solution is presented as Gaussian sum with explicitly computed specific weights, means and variances. However, a number of statistics grows geometrically in time, which restricts duration of online estimation by number of time steps. In contrast to that, the approach proposed in this paper runs online and does not cause increasing the statistics in time.

Other methods closely related to that described in this paper are, for example, iterative techniques nicely presented in Doucet and Andrieu (2001) and the mixture Kalman filter (Chen & Liu, 2000). A part concerned with the estimation of discrete multinomial state is also close to HMM theory (Beal, Ghahramani, & Rasmussen, 2002). However, the algorithms mentioned run mostly offline and are supported by Monte Carlo computations. The presented paper aims at online state estimation and exploitation of explicit solutions using numerical procedures only in that parts, which cannot be computed analytically.

Comparison with one of the main counterparts – the interactive multiple model (IMM) filter (Bar-Shalom et al., 2002) – is presented.

Acknowledgements

The research was partially supported by projects MŠMT 1M0572 and TAČR TA01030123 and general agreement of ÚTIA and Škoda Auto No. ENS/2009/UTIA.

References

Bar-Shalom, Y., Kirubarajan, T. & Li, X.-R. (2002). *Estimation with applications to tracking and navigation*. New York, NY, USA: Wiley.
Beal, M. J., Ghahramani, Z. & Rasmussen, C. E. (2002). The infinite hidden Markov model. In *Advances in neural information processing systems, vol. 14*.
Chen, R. & Liu, J. S. (2000). Mixture kalman filters. *Journal of Royal Statistical Society B*, *62*, 493–508.
Doucet, A. & Andrieu, C. (2001). Iterative algorithms for state estimation of jump Markov linear systems. *IEEE Transactions on Signal Processing*, *49*(6), 1216–1227.
Elliot, F. D. R. J. & Sworder, D. (1996). Exact hybrid filters in discrete time. *IEEE Transactions on Automatic Control*, *41*(12), 1807–1810.
Grewal, M. & Andrews, A. (2001). *Kalman filtering: Theory and practice using MATLAB* (2nd ed.). Wiley.
Kadlec, P., Gabrys, B. & Strandt, S. (2009). Data-driven soft sensors in the process industry. *Computers and Chemical Engineering*, *33*, 795–814.
Kárný, M., Böhm, J., Guy, T. V., Jirsa, L., Nagy, I., Nedoma, P., et al. (2005). *Optimized Bayesian dynamic advising: Theory and algorithms*. London: Springer.
Kerridge, D. (1961). Inaccuracy and inference. *Journal of Royal Statistical Society*, *B 23*, 284–294.
Kullback, S. & Leibler, R. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, *22*, 79–87.
Lin, B., Recke, B., Renaudat, P., Knudsen, J. & Jörgensen, S. B. (2005). A systematic approach for soft sensor development. *Computers and Chemical Engineering*, *5*, 1147–1152.
Park, S. & Han, C. (2000). A nonlinear soft sensor based on multivariate smoothing procedure for quality estimation in distillation columns. *Computers and Chemical Engineering*, *24*, 871–877.
Peterka, V. (1981). Bayesian system identification. In P. Eykhoff (Ed.), *Trends and progress in system identification.* (pp. 239–304). Oxford: Pergamon Press.
Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N. & Yin, K. (2003). A review of process fault detection and diagnosis. Part III. Process history based methods. *Computers and Chemical Engineering*, *27*, 327–346.