Probabilistic Approach to Analysis of Death Traffic Accidents

Evgenia Suzdaleva and Ivan Nagy

***Abstract*— The paper is devoted to analysis of data related to traffic accidents at one of the roads in Czech Republic. The data sets are available as discrete-valued variables providing results of traffic accident (with death or not) as well as conditions under which the accident has happened (weather, visibility, speed etc). Situation of a traffic accident is modeled within state-space framework. Estimation of accidents results is proposed with the help of Bayesian filtering based on determined probabilities.**

I. INTRODUCTION

The paper deals with analysis of data sets concerned with death traffic accidents on the road II/114 on Czech Republic. Everyone knows that in out time of progress and high speed powerful cars the death accident statistics remains unfavorable and offensive: it is enough to see the everyday news. Modeling of traffic accidents and conditions under which they happened is important task to predict deathrate on certain roads under certain driving modes, speed, visibility etc. It gives a chance a driver to better regulate controllable variables such as speed and driving modes via advices of intelligent built-in car computers. This motivated us to propose our solution presented at this paper.

One of the accompanied problems of modeling the traffic accidents is a lack of informative data and small amount of data sets. The available data sets contain measurable variables, mostly of a seasonable character, which describe various conditions of the accident. The result of the accident plays a role of a modeled variable. Surely, there are many other unmeasured variables, influencing it. The data sets can be explored via various approaches to extract the information they can bring and understand relationships between variables. However, a choice of accurate model of the traffic accident is a difficult and ambiguous task, which requires an extra-use of prior expert knowledge. Previously, the available data sets have been modeled with the help of logistic regression [1]. This is the most simple way of modeling, remained, however, in a static way. Another way, proposed in this paper, is to be aware of the unmeasured variables, or even to reduce the number of variables in the data vector. The unmeasured and omitted variables, which are mostly periodical with the period one year, cause dynamics of the modeled variable. Thus, the dependence on some external variables can be substituted by considering the dynamic (state-space) model of the variable monitored. The present paper proposes to set the traffic accident as a state variable and use the state-space model [2]. The state-space model is considered for discrete-valued variables with Bernoulli distribution. Relationships between the variables are supposed via probabilities of a certain event, defined by traffic experts (or estimated offline). Physically, it can be interpreted as in the following example. In winter there are much more mortal accidents then in summer. So, if the last accident was mortal (and the season was not known) it is probable that it happened during winter and that the winter still lasts. It means, that the mortal accident is now more probable, too. This interpretation explains usage of probabilities in the discrete state-space model. The time moments used in the dynamic model are irregular ones, determined by a sequence of the traffic accidents. The variables from the available data sets are exploited as a discrete output and a discrete input for the considered model. At the present paper, the estimation of a discrete result of traffic accident within a given set of possibles values (either with death or injury or only with slight material damage) is proposed as an analytically tractable solution of Bayesian filtering [3]. The proposed solution does not require high computational cost (as, for example, particle filters [4], which can be possibly used for this aim) and thus it can be exploited in real conditions with available expert knowledge.

A layout of the paper is as follows. Section II provides basic facts about the models used as well as Bayesian filtering. Section III is devoted to the filtering with discrete-valued variables and presents the algorithm. Section IV describes specialization of the proposed technique to analysis of the death traffic accidents and provides results.

II. PRELIMINARIES

The state-space model to be used in the paper is provided with the help of the following conditional probability (density) functions (p(d)fs). The *observation model*, specified by pdf

$$f\left(y_t|u_t, x_t\right), \quad (1)$$

relates the system output y_t to the system input u_t and the unobserved system state x_t at discrete time moments $t \in t^* \equiv \{0, \dots, \mathring{t}\}$, where $\mathring{t}$ is the cardinality of the set t^* and $\equiv$ means equivalence. The *state evolution model*

$$f\left(x_{t+1}|u_t, x_t\right), \quad (2)$$

describes the evolution of the system state x_t. The estimation of the finite-dimensional system state calls for application of

This work was supported by project MŠMT 1M0572.

E. Suzdaleva is with the Departmentt 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 suzdalev@utia.cas.cz

Ivan Nagy is with Faculty of Transportation Sciences, Czech Technical University, Na Florenci 25, 11000 Prague, Czech Republic nagy@utia.cas.cz

Bayesian filtering. Bayesian filtering, estimating the system state, includes the following coupled formulas.
Data updating

$$f\left(x_t\middle|d^t\right) = \frac{f\left(y_t\middle|u_t,x_t\right)f\left(x_t\middle|d^{t-1}\right)}{\int f\left(y_t\middle|u_t,x_t\right)f\left(x_t\middle|d^{t-1}\right)\,dx_t}, \propto f\left(y_t\middle|u_t,x_t\right)f\left(x_t\middle|d^{t-1}\right), \quad (3)$$

($\propto$ means proportionality) incorporates the experience contained in the data d^t, where $d^t = (d_0,\ldots,d_{\mathring{t}})$ and $d_t \equiv (y_t,u_t)$.
Time updating

$$f\left(x_{t+1}\middle|d^t\right) = \int f\left(x_{t+1}\middle|u_t,x_t\right)f\left(x_t\middle|d^t\right)\,dx_t, \quad (4)$$

fulfills the state prediction. The filtering does not depend on the control strategy $\{f(u_t|d^{t-1})\}_{t\in t^*}$ but on the generated inputs only. The prior pdf $f(x_0)$, which expresses the subjective prior knowledge on the state x_0, starts the recursions.

Bayesian filtering (3)-(4) can be analytically solved with linear Gaussian models. In that case the solution coincides with Kalman filter [5]. With discrete variables, the hidden Markov models [6], [7], [8] as well as sampling-based solutions are often exploited. Due to limited computational cost in the present application domain the particle filters are hardly used. The paper proposes a non-approximative solution for a relatively simple case of the filtering with Bernoulli models.

III. Non-approximative filtering with discrete-valued state

The considered variables y_t, x_t and u_t are of a discrete-valued nature. It means that one must investigate application of the filtering (3)-(4) to models with discrete distributions. When speaking about the state, one must consider the state vector $x_t \equiv [x_{1;t},\ldots,x_{\mathring{x};t}]'$ with finite, preferably small $\mathring{x}$, where each entry $x_{i;t}$ with $i = \{1,\ldots,\mathring{x}\}$ has a set of possible values $\{s_1,s_2,\ldots,s_n\}$ with finite number n. The estimation of the state vector with discrete entries is simplified via a special mapping of the multivariate state to a scalar one. The mapping is proposed via a determination of the new set of possible values for the scalar discrete state. The new set is constructed so that each possible combination of values of all entries $x_{i;t}$, i.e. $\{(x_{1;t}=s_1,x_{2;t}=s_1,\ldots,x_{\mathring{x};t}=s_1),(x_{1;t}=s_2,x_{2;t}=s_1,\ldots,x_{\mathring{x};t}=s_1),\ldots,(x_{1;t}=s_n,x_{2;t}=s_n,\ldots,x_{\mathring{x};t}=s_n)\}$, is denoted as the new scalar value, belonging to this set, i.e. $\{k_1,k_2,\ldots,k_N\}$, N is a number of combinations[1]. The mapping provides necessary reduction of dimension of the state. The similar reducing of the dimension is applied to the multivariate output and input. Let's assume, that the scalar variables (either due to the mapping or naturally) with the finite set of possible discrete values have to be considered for the filtering. Generally, discrete multinomial distribution could have been used for the models (1)-(2) with these variables. For the sake of simplicity, a number of possible values in the set is restricted by two, which leads to Bernoulli distribution.

Let's consider the observation model (1) described by the Bernoulli distribution, shown in Table I, where α with corresponding indices denotes a probability (assumed to be known) of taking the possible values by the output, conditioned on values of the state and input. The set of discrete values for all the variables is given as $\{k_1,k_2\}$, and the probability $\alpha_{2|ij}$ is always defined as $(1-\alpha_{1|ij})$.

TABLE I
Bernoulli Observation Model

	$y_t = k_1$	$y_t = k_2$
$u_t = k_1, x_t = k_1$	$\alpha_{1\|11}$	$\alpha_{2\|11}$
$u_t = k_2, x_t = k_1$	$\alpha_{1\|21}$	$\alpha_{2\|21}$
$u_t = k_1, x_t = k_2$	$\alpha_{1\|12}$	$\alpha_{2\|12}$
$u_t = k_2, x_t = k_2$	$\alpha_{1\|22}$	$\alpha_{2\|22}$

The Bernoulli distribution from Table I written with the help of Kronecker delta is presented in the following product form

$$f\left(y_t\middle|u_t,x_t\right) = \prod_{u_t,x_t\in\{k_1,k_2\}} \alpha_{1|u_tx_t}^{\delta(y_t,k_1)}\alpha_{2|u_tx_t}^{\delta(y_t,k_2)}, \quad (5)$$

where Kronecker delta expresses a choice of an occurred value from the possible ones. Similarly, the state evolution model (2) is related to Bernoulli distribution, provided in Table II, where respective β denote a known probability of taking the possible values of the state, conditioned on its previous values and on the input, and $\beta_{2|ij} = (1-\beta_{1|ij})$. The

TABLE II
Bernoulli State Evolution Model

	$x_{t+1} = k_1$	$x_{t+1} = k_2$
$u_t = k_1, x_t = k_1$	$\beta_{1\|11}$	$\beta_{2\|11}$
$u_t = k_2, x_t = k_1$	$\beta_{1\|21}$	$\beta_{2\|21}$
$u_t = k_1, x_t = k_2$	$\beta_{1\|12}$	$\beta_{2\|12}$
$u_t = k_2, x_t = k_2$	$\beta_{1\|22}$	$\beta_{2\|22}$

product form of the distribution from Table II is as follows.

$$f\left(x_{t+1}\middle|u_t,x_t\right) = \prod_{u_t,x_t\in\{k_1,k_2\}} \beta_{1|u_tx_t}^{\delta(x_{t+1},k_1)}\beta_{2|u_tx_t}^{\delta(x_{t+1},k_2)}. \quad (6)$$

The prior probabilities for the initial discrete state are chosen as $p_{1(t)}$ for value $x_t = k_1$ and $p_{2(t)} = (1-p_{1(t)})$ for $x_t = k_2$. Thus, the form of the prior Bernoulli distribution is defined as

$$f\left(x_t\middle|d^{t-1}\right) = p_{1(t)}^{\delta(x_t,k_1)}(1-p_{1(t)})^{\delta(x_t,k_2)}. \quad (7)$$

The estimation of the discrete state is proposed as the direct application of Bayesian filtering (3)-(4) to the Bernoulli state-space model (5)-(6) with incorporation of Bernoulli prior (7). According to the mentioned relations, formula (3) with the substituted Bernoulli distributions (5) and (7) takes the following form, providing the updating of the state estimate

[1]For example, for $x_t \equiv [x_{1;t},x_{2;t}]'$, with $x_{1;t}\in\{s_1,s_2\}$ and $x_{2;t}\in\{s_1,s_2\}$, the possible values $\{(s_1,s_1),(s_2,s_1),(s_1,s_2),(s_2,s_2)\}$ can be denoted as the new set $\{k_1,k_2,k_3,k_4\}$.

by actual measurements

$$f\left(x_t\middle|d^t\right) = \frac{\prod_{u_t,x_t\in\{k_1,k_2\}}\alpha_{1|u_tx_t}^{\delta(y_t,k_1)}\alpha_{2|u_tx_t}^{\delta(y_t,k_2)}}{\sum_{x_t\in\{k_1,k_2\}}\prod_{u_t,x_t\in\{k_1,k_2\}}\alpha_{1|u_tx_t}^{\delta(y_t,k_1)}} \quad (8)$$

$$\frac{p_{1(t)}^{\delta(x_t,k_1)}(1-p_{1(t)})^{\delta(x_t,k_2)}}{\alpha_{2|u_tx_t}^{\delta(y_t,k_2)}p_{1(t)}^{\delta(x_t,k_1)}(1-p_{1(t)})^{\delta(x_t,k_2)}}$$

$$= \bar{p}_{1(t)}^{\delta(x_t,k_1)}\bar{p}_{2(t)}^{\delta(x_t,k_2)}, \quad (9)$$

where integration in the denominator is replaced by regular summation. The probabilities to be substituted in the data updating (8) are chosen from Table I according to the actual values of the output and the input. $\bar{p}(\cdot)$ in (9) denotes the intermediate results of the filtering (data-updated probabilities).

The Bayesian time updating (4) with Bernoulli distribution (6) and the intermediate result (9) takes the following form

$$f\left(x_{t+1}\middle|d^t\right) = \sum_{x_t\in\{k_1,k_2\}}\prod_{u_t,x_t\in\{k_1,k_2\}}\beta_{1|u_tx_t}^{\delta(x_{t+1},k_1)} \beta_{2|u_tx_t}^{\delta(x_{t+1},k_2)}\bar{p}_{1(t)}^{\delta(x_t,k_1)}\bar{p}_{2(t)}^{\delta(x_t,k_2)}, \quad (10)$$

which provides the resulting state estimate as the following Bernoulli distribution

$$f\left(x_{t+1}\middle|d^t\right) = p_{1(t+1)}^{\delta(x_{t+1},k_1)}(1-p_{1(t+1)})^{\delta(x_{t+1},k_2)}, \quad (11)$$

with the updated probability of value $x_{t+1} = k_1$

$$p_{1(t+1)} = \sum_{x_t\in\{k_1,k_2\}}\prod_{u_t,x_t\in\{k_1,k_2\}}\beta_{1|u_tx_t}^{\delta(x_{t+1},k_1)}\bar{p}_{1(t)}^{\delta(x_t,k_1)}\bar{p}_{2(t)}^{\delta(x_t,k_2)} = \beta_{1|u_t1}\bar{p}_{1(t)} + \beta_{1|u_t2}\bar{p}_{2(t)}, \quad (12)$$

calculated according to the known values of the input and substitution of the corresponding probabilities from Table II. The probability of value k_2 is obtained as $p_{2(t+1)} = (1 - p_{1(t+1)})$, or can be calculated directly:

$$p_{2(t+1)} = \sum_{x_t\in\{k_1,k_2\}}\prod_{u_t,x_t\in\{k_1,k_2\}}\beta_{2|u_tx_t}^{\delta(x_{t+1},k_2)}\bar{p}_{1(t)}^{\delta(x_t,k_1)}\bar{p}_{2(t)}^{\delta(x_t,k_2)} = \beta_{2|u_t1}\bar{p}_{1(t)} + \beta_{2|u_t2}\bar{p}_{2(t)}. \quad (13)$$

Relation (13) is obtained similarly according to the input values and substitution of the probabilities from Table II. The resulting Bernoulli distribution (11) is taken as the prior one to be incorporated into the next step of the discrete state estimation (8) with actual available measurements.

Algorithm

1) *Set prior $p_{2(t=0)}$ for value $x_t = k_2$*
2) *Get actual data y_t, u_t*
3) *Choice from Table I.*
   If $y_t = k_2$ & $u_t = k_2$, *then* $\alpha = \alpha_{2|22}, g = \alpha_{2|21}$, *where g is an auxiliary variable for normalization, conditioned on* $x_t = k_1$,
   else if $y_t = k_2$ & $u_t = k_1, \alpha = \alpha_{2|12}, g = \alpha_{2|11}$,
   else if $y_t = k_1$ & $u_t = k_2, \alpha = \alpha_{1|22}, g = \alpha_{1|21}$,
   else if $y_t = k_1$ & $u_t = k_1, \alpha = \alpha_{1|12}, g = \alpha_{1|11}$,
   end
4) *Data updating.*
   $\bar{p}_{2(t)} = \frac{\alpha\cdot p_{2(t)}}{\alpha\cdot p_{2(t)}+g\cdot(1-p_{2(t)})}$, $\bar{p}_{1(t)} = 1 - \bar{p}_{2(t)}$
5) *Choice from Table II.*
   If $u_t = k_1$, *then* $\beta = \beta_{2|12}, d = \beta_{2|11}$, *where d is an auxiliary variable, conditioned on* $x_t = k_1$,
   else if $u_t = k_2, \beta = \beta_{2|22}, g = \beta_{2|21}$,
   end
6) *Time updating.*
   $p_{2(t)} = \beta\cdot\bar{p}_{2(t)} + d\cdot\bar{p}_{1(t)}$, $p_{1(t)} = 1 - p_{2(t)}$. *Go to* 2.

The algorithm is presented for the Bernoulli distribution, but it can be easily extended up to the multinomial one. However, in practice the multinomial distribution complicates the proposed filtering by setting of the larger number of the probabilities for the observation and state-evolution models.

IV. Death traffic accidents analysis and results of experiments

The available data sources represent the evidence about accidents on the road II/114 in the Czech republic. The modeled variable x_t relates to a result of the traffic accident so that $x_t \in \{1, 0\}$, where 1 = accident with material damage, 0 = accident caused death or injury. The measurements D_t contain the following information:

- $D_{1;t}$ is a daytime, where 1 – day, 2 – dawn, dusk, 3 – night;
- $D_{2;t}$ is visibility, where 1 – clear weather, 2 – fog, 3 – rain, 4 – snow;
- $D_{3;t}$ is a speed with 1 – normal, 2 – high;
- $D_{4;t}$ is a cause of accident, where 1 – high speed, 2 – wrong driving, 3 – wrong overtaking, 4 – other;
- $D_{5;t}$ is a type of accident: 1 – danger, 2 – crash between cars, 3 – crash with a fixed object, 4 – collision with an animal.

The mentioned data sources have been analyzed with the help of various approaches. The understanding if the variables bring some information sufficient for modeling was the extremely difficult task due to small capacity of the available data sets. The idea to use the state-space model with the described probabilistic approach came after the attempt to make prediction via logistic regression. The results of exploitation of logistic regression is shown in Fig. 1. The quality of the estimation shown significant dependence on prior knowledge. It can be seen that the obtained results can be improved. The exploitation of the proposed algorithm of the discrete-valued state filtering did not required high computational cost and gave a possibility to incorporate expert knowledge. That's the way it was decided to test the algorithm on the available data. The data sets have been rescaled until two possible values of the modeled variables so that to be simplified and also adapted to Bernoulli model. Here, two experiments are demonstrated: one with scalar variables and another with reduced dimension of the variables via the proposed mapping. The last experiment was obviously complicated from the point of view of expert setting of probabilities for models (5)-(6). The data sets that are the most informative for filtering have been chosen for experiments.

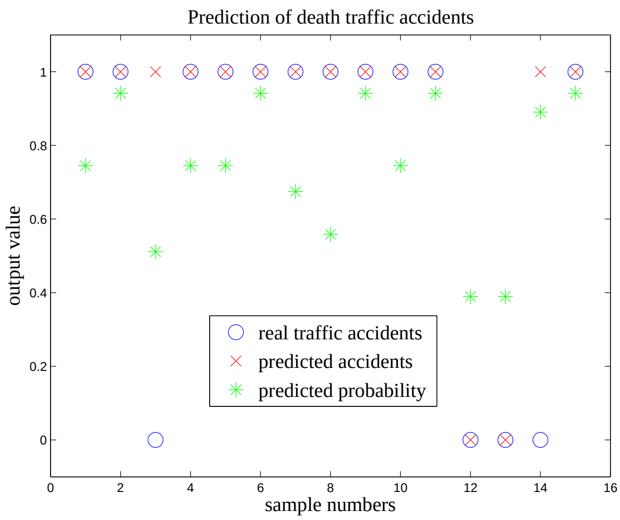

Fig. 1. Prediction of traffic accidents with the logistic regression

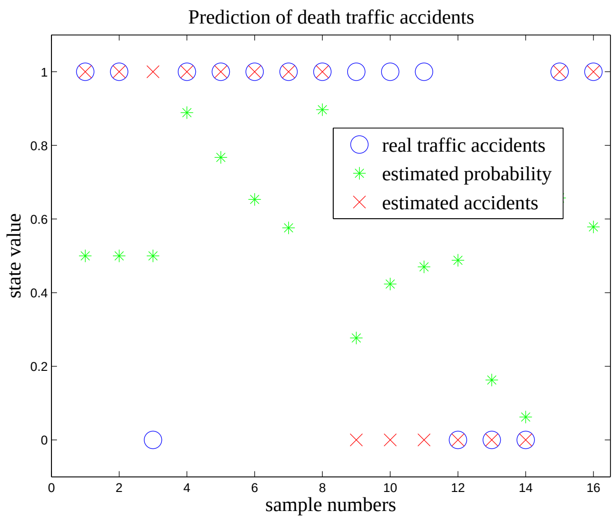

Fig. 2. Filtering of death traffic accidents with scalar variables

A. Experiment 1

The available values of the variable $D_{2;t}$ were identified with the system output according to (5). The visibility $D_{2;t} \in \{1, 0\}$ was rescaled so that 1 means clear visibility (with clear weather), and 0 denotes the worse visibility with fog, rain or snow. Variable $D_{3;t} \in \{1, 0\}$ expressing the measured speed has been identified with the input in (5) so that 1 means the normal speed, while 0 denotes the high speed. The prior probability of value $x_t = 0$ (i.e. the death accident) is chosen to be equal 0.5. Parameters (probabilities) given by traffic experts for observation model (5) (for prediction of measurements) and for state evolution model (6) are shown in Tables III-IV respectively. The results shown at Fig. 2 demonstrate comparison of real and estimated results of the accidents. The results of the filtering strongly depend on expert knowledge given via parameters of observation and state evolution models (5)-(6). This simple experiment was obviously made for testing of the proposed algorithm. Improvement of the obtained results is expected to be reached with the help of offline estimation of parameters, which is planned in later work.

TABLE III
PROBABILITIES FOR OBSERVATION MODEL

	$D_{2;t} = 0$	$D_{2;t} = 1$
$D_{3;t} = 1, x_t = 1$	0.68	0.32
$D_{3;t} = 0, x_t = 1$	0.49	0.51
$D_{3;t} = 1, x_t = 0$	0.04	0.96
$D_{3;t} = 0, x_t = 0$	0.45	0.55

TABLE IV
PROBABILITIES FOR STATE EVOLUTION MODEL

	$x_{t+1} = 0$	$x_{t+1} = 1$
$D_{3;t} = 1, x_t = 1$	0.08	0.92
$D_{3;t} = 0, x_t = 1$	0.69	0.31
$D_{3;t} = 1, x_t = 0$	0.64	0.36
$D_{3;t} = 0, x_t = 0$	0.97	0.03

B. Experiment 2

For this experiment the cause of the accident $D_{4;t}$ was identified with the output in the model (5). It was rescaled so that 1 means the high speed caused the accident and 0 comprise all reasons concerned with a wrong driving. This rescaling has been chosen according to presence of more informative data in the data sets. The input in models (5)-(6) was identified with a vector $D_t^r \equiv [D_{2;t}; \quad D_{3;t}]$. The mapping proposed in Section III has been applied to the vector D_t^r to reduce dimension until a scalar. It results in the following set of possible values for the reduced variable $D_t^r \in \{D1, D2, D3, D4\}$, where $D1 \equiv [1;\ 1]$, $D2 \equiv [0;\ 1]$, $D3 \equiv [1;\ 0]$, $D4 \equiv [0;\ 0]$. Here, setting of probabilities for models (5)-(6) was more complicated task than in the previous experiment. The parameters for (5)-(6) provided by experts are shown in Tables V-VI respectively. The results of the estimation are shown at Fig. 3.

TABLE V
PROBABILITIES FOR OBSERVATION MODEL

	$D_{4;t} = 0$	$D_{4;t} = 1$
$D_t^r = D1, x_t = 1$	0.87	0.13
$D_t^r = D2, x_t = 1$	0.73	0.27
$D_t^r = D3, x_t = 1$	0.04	0.96
$D_t^r = D4, x_t = 1$	0.11	0.89
$D_t^r = D1, x_t = 0$	0.96	0.04
$D_t^r = D2, x_t = 0$	0.72	0.28
$D_t^r = D3, x_t = 0$	0.18	0.82
$D_t^r = D4, x_t = 0$	0.14	0.86

V. CONCLUSIONS AND FUTURE WORKS

The paper describes the approach to the analysis of data sources representing the evidence about death traffic accidents on one of the Czech roads. The aim of the paper was to test the proposed filtering algorithm on the traffic

TABLE VI

PROBABILITIES FOR STATE EVOLUTION MODEL

	$x_{t+1} = 0$	$x_{t+1} = 1$
$D_t^r = D1, x_t = 1$	0.04	0.96
$D_t^r = D2, x_t = 1$	0.21	0.79
$D_t^r = D3, x_t = 1$	0.83	0.17
$D_t^r = D4, x_t = 1$	0.82	0.18
$D_t^r = D1, x_t = 0$	0.96	0.04
$D_t^r = D2, x_t = 0$	0.94	0.06
$D_t^r = D3, x_t = 0$	0.12	0.88
$D_t^r = D4, x_t = 0$	0.7	0.3

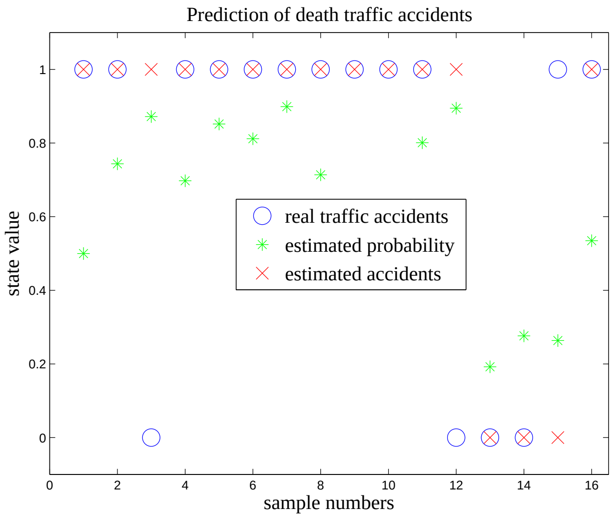

Fig. 3. Filtering of death traffic accidents with mapping

accidents data sets. To conclude the paper, one can say that the testing opens weak points of the algorithm. The algorithm is very sensitive to the setting of probabilities for the model, provided by traffic experts. It should be noted that better results of prediction of *death* accidents are expected with the help of offline estimation of parameters. Thus, future work will be concerned with the parameter estimation for improvements of the filtering with the discrete state-space model.

A really important remark to the present paper is the point that the authors neither search nor invent the new optimal filtering algorithm for discrete models. The described filtering is one of items aiming at filtering with mixed (continuous and discrete) models. It means that the global objective of the research is to find solution for the joint filtering of the mixed-type variables (more information is available in [9]).

VI. ACKNOWLEDGMENTS

This work was supported by project MŠMT 1M0572.

REFERENCES

[1] D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression*, Wiley-Interscience, 2001.

[2] V. Peterka, "Bayesian system identification," in *Trends and Progress in System Identification*, P. Eykhoff, Ed., pp. 239–304. Pergamon Press, Oxford, 1981.

[3] E. Suzdaleva, I. Nagy, and L. Pavelková, "Bayesian filtering with discrete-valued state", in *Proceedings of IEEE workshop on Statistical Signal Processing*, Cardiff, Wales, UK, Aug.31-Sept.3 2009.

[4] V. Šmídl and A. Quinn, "Variational bayesian filtering," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 5020–5030, 2008.

[5] M.S. Grewal and A.P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB. 2nd edition*, Wiley, 2001.

[6] L. Shue, B.D.O. Anderson, S. Dey, "Exponential stability of filters and smoothers for hidden Markov models," *IEEE Transactions on Signal Processing*, vol.46(8), pp. 2180 - 2194, Aug. 1998.

[7] Chun Yang, "On discrete hidden markov state estimation," in *Proceedings of the American Control conference*, Seattle, WA, USA, June 21-23 1995, vol. 1, pp. 12–13.

[8] S. Di Cairano, K. H. Johansson, A. Bemporad, and R. M. Murray, *Hybrid Systems: Computation and Control*, vol. 4981/2008, chapter Discrete and Hybrid Stochastic State Estimation Algorithms for Networked Control Systems, pp. 144–157, Springer, 2008.

[9] E. Suzdaleva, "Filtering with mixed continuous and discrete states: special case," Tech. Rep. 2246, ÚTIA AV ČR, Praha, January 2009, Draft of paper. http://library.utia.cas.cz/separaty/2009/AS/suzdaleva-filtering with mixed continuous and discrete states special case.pdf