Particle Filters For Max Plus Systems

Renato Markele Ferreira Cândido, Rafael Santos Mendes (Orientador)

Department of Computer Engineering and Industrial Automation (DCA)
Faculty of Electrical and Computer Engineering (FEEC)
University of Campinas (Unicamp)
Av. Albert Einstein, 400 - Campinas, SP, Brazil - CEP: 13083-852
{rmarkele,rafael}@dca.fee.unicamp.br

Abstract – The main objective of this article is to synthesize a particle filter algorithm for max-plus systems. It is presented a brief introduction to the max-plus approach for Discrete Event Systems. Next, the fundamentals of the particle filter and the way in which they can be applied to max-plus systems are presented. It leads to the algorithm for particles filtering. Lastly, an example is given. The results shows the accuracy of the method and the improvements in comparison with the deterministic observer.

Keywords – Stochastic filtering; Discrete event systems; Petri nets

1. Introduction
In this paper one attempts to estimate sequentially the state of a Discrete Event System (DES) in a noise context, where the state is not fully observable, by using a sub-optimal Bayesian filter: The Particle Filters (PFs). These filters performs a Sequential Monte Carlo estimation to produce an approximation of a probability distribution based on a set of samples (particles) with associated weights.

The Discrete Event (Dynamic) Systems (DEDS) are discrete-state systems whose state evolution depends entirely on the occurrence of asynchronous discrete events over time [2]. Among the existing models for treating these systems stand out the approach based in the Max-Plus Algebra [1], which is the model adopted in this work to obtain the equation that describes the evolution of the states and the equation that relates the states with the output.

2. Petri Nets and Max-Plus Algebra
Petri Nets describe DEDS pictorially and can be viewed as a bipartite graph. A timed event graph (TEG) is a Petri net in which all places have exactly one upstream and one downstream transition. In this paper, the $p$-timed Petri nets are used. In these networks each place is associated with a minimum time of permanency for the tokens.

Consider the Fig. 1. For each transition $T_i$, one can associate a non-decreasing sequence formed by the variables $x_i(k)$, $k = 1, 2, ..., $ called daters, containing the $k$-th instant of firing of the transition $T_i$. Assuming that the sequences of firing associated with the input transitions are known it is possible to determine the sequences of firing

\[
x_1(k) = \text{max}\{10 + x_1(k - 1); 7 + x_2(k - 1); \ u(k)\}
\]

\[
x_2(k) = \text{max}\{13 + x_1(k - 1); 10 + x_2(k - 1); \ 3 + u(k)\}
\]

\[
z_1(k) = 15 + x_1(k)
\]

\[
z_2(k) = 12 + x_2(k)
\]

which, in matrix form, is:

\[
x_k = A \otimes x_{k-1} \oplus B \otimes u_k \quad (1a)
\]

\[
z_k = C \otimes x_k \quad (1b)
\]

Figure 1. A timed events graph
Therefore, all transitions of the timed event graph were described by a recursive linear system of equations. The set $\mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ is called Max-Plus and noted by $\mathbb{R}_{\text{max}}$ [1].

This paper considers discrete event systems that can be described by equations (1).

3. Particles Filters Applied to Max Plus Systems

The Particle Filters are suboptimal filters that perform a Sequential Monte Carlo (SMC) estimation based on particle representation of probability densities. The following is a brief description of the Monte-Carlo method [6].

Let $I = \int g(x)dx$ be an integral that we wish to solve and suppose $g(x)$ can be factored as $g(x) = f(x) \cdot \pi(x)$ in such way that $\pi(x)$ is a probability density. It is generally impossible to sample from $\pi(x)$ directly. So, assuming that it is possible to drawn $N \gg 1$ samples $\{x^i, i = 1, ..., N\}$ distributed according to $g(x)$, similar to $\pi(x)$, the Monte Carlo estimate integral is

$$I_N = \frac{1}{N} \sum_{i=1}^{N} f(x^i) \cdot \omega(x^i),$$

where $\omega(x^i) = \frac{\pi(x)}{q(x)}$. This process is called importance sampling.

Now, let $\mathbf{X}_k = \{x_j, j = 0, ..., k\}$ be the sequence of the first $k$ firing times of all states of a system described by equations (1). The joint posterior density at $k$ is denoted by $p(\mathbf{X}_k|\mathbf{Z}_k)$ and its marginal is $p(x_k|\mathbf{Z}_k)$. Assuming the existence of a set of particles (or samples) denoted by $\{\mathbf{X}_k^i, i = 1, ..., N\}$ and their respective weights $\{\omega_k^i, i = 1, ..., N\}$, the joint posterior at $k$ can be approximated as:

$$p(\mathbf{X}_k|\mathbf{Z}_k) \approx \sum_{i=1}^{N} \omega_k^i \delta(\mathbf{X}_k - \mathbf{X}_k^i)$$

(2)

If the particles $\mathbf{X}_k^i$ were drawn from an importance density $q(\mathbf{X}_k|\mathbf{Z}_k)$ then:

$$\omega_k^i \propto \frac{p(\mathbf{X}_k^i|\mathbf{Z}_k)}{q(\mathbf{X}_k^i|\mathbf{Z}_k)}$$

(3)

A probability density $q(x)$ (usually referred to as importance weights) is similar to $\pi(x)$ if $\forall x : \pi(x) > 0 \Rightarrow q(x) > 0$.

Therefore, it is desirable to find a recursive equation in such way that, having an approximation of density $p(\mathbf{X}_{k-1}|\mathbf{Z}_{k-1})$ composed of a set of particles at $k - 1$, an approximation of $p(\mathbf{X}_k|\mathbf{Z}_k)$ can be calculated with a new set of particles as soon as a new measurement $z_k$ becomes available. It has been shown in [6] that the desired equation is given by:

$$\omega_k^i = \omega_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{k-1}^i, z_k)}$$

(4)

In this paper the following suboptimal choice for the importance density has been adopted:

$$q(x_k|x_{k-1}^i, z_k) = p(x_k|x_{k-1}^i).$$

(5)

This choice allows the use of the equation (1a) to augment each existing particle $\mathbf{X}_{k-1}^i$ to $\mathbf{X}_k^i$.

By substituting (5) in (4) one obtains:

$$\omega_k^i = \omega_{k-1}^i \cdot V(x_k, z_k),$$

(7)

where $V(x, z)$ is the likelihood function given by [3]:

$$V(x, z) = \prod_{i=1}^{q} \prod_{j=1}^{n} p_{ij}(z_i - x_j) \prod_{k=1, k \neq j}^{n} F_{ik}(z_i - x_k)$$

(8)

where, it is assumed that the matrix elements of $C$ (Eq. (1)), noted by $c_{ij}$, are independent random variables uniformly distributed with cumulative probability and probability density functions represented by $F_{ij}(\cdot)$ and $p_{ij}(\cdot)$ respectively.

The estimation of $x_k$ is therefore given by:

$$\hat{x}_k = \sum_{i=1}^{N} x_k^i \omega_k^i.$$
occurs when, after a certain number of recursive steps, many particles get very low weights. This phenomenon is impossible to avoid but a process of resampling [6] is a strategy to overcome degeneracy of samples. In this process the particles with low weights are eliminated and the particles with non-negligible weights are cloned proportionally to their respective weights.

In general, the update of the weights is done by the likelihood function. However a problem occurs when there is a deterministic relationship between the output and the state. For example, assume that in a given system the relationship between the output and the state is given by (10).

\[
z(k) = \begin{bmatrix} 2 \mu(2,3) \end{bmatrix} \otimes \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}
\]

where \( \mu(\bar{x}, \bar{x}) \) represents a uniform random variable distributed between \( \bar{x} \) and \( \bar{x} \).

The equation (10) can be seen as \( z(k) = C x(k) \) with \( C = \begin{bmatrix} 2 \mu(2,3) \end{bmatrix} \). The cumulative probability and probability density functions of element \( c_{11} \) are given respectively by:

\[
F_{11}(\tau) = \begin{cases} 
0 & \text{if } \tau < 2 \\
1 & \text{if } \tau \geq 2 
\end{cases}
\]

\[
p_{11}(\tau) = \delta(\tau - 2)
\]

where \( \delta(\tau) \) is the impulse function.

Consider that at a certain iteration \( k \), the measurement \( z(k) = 5 \) is acquired. Consider also that after the process of predicting the state, the following set of particles approximating the prior density \( p(x_k|x_{k-1}) \) was available:

\[
x_k = \begin{bmatrix} 2.05 & 0.60 & 2.80 & 1.80 & 2.95 & 3.05 & 2.95 \\
1.80 & 2.80 & 1.70 & 2.20 & 2.80 & 1.20 & 0.30 \end{bmatrix}
\]

where each column represents a particle \( x_k^i = [x_1^i(k) \ x_2^i(k)]^T \). Using (8) one can calculate the likelihood of each particle \( x_k^i \):

\[
V(x_k^i, z) = 0,
\]

\[
V(x_2^i, z) = 1, V(x_3^i, z) = 0, V(x_4^i, z) = 1, V(x_5^i, z) = 1,
\]

\[
V(x_6^i, z) = 0 \text{ and } V(x_7^i, z) = 0.
\]

For this simple example, is easy to conclude that the posterior possible values for the state \( [x_1(k) \ x_2(k)] \), given the measure \( z(k) = 5 \), belong to region \( R_1 \cup R_2 \), where \( R_1 = \{ [x_1(k) \ x_2(k)] \in \mathbb{R}^2 : x_1(k) = 3 \text{ and } x_2(k) < 2 \} \) and \( R_2 = \{ [x_1(k) \ x_2(k)] \in \mathbb{R}^2 : x_1(k) \leq 3 \text{ and } 2 \leq x_2(k) \leq 3 \} \), depicted in Fig. 2. The region \( R_1 \) is a half line that corresponds to an event with null probability measure in the prior probability space, i.e. the probability of a particle be within \( R_1 \) is zero. In the particular case where all the elements of \( C \) are deterministic this would lead to a step of updating of the weights where all weights would become zero.

It should be noted however that, if the measure \( z(k) = 5 \) occurs then the state belongs to the feasible region \( R_1 \cup R_2 \). So, the prior distribution of the particles must be rearranged in order to undertake this fact. This rearrangement of the particles can be done by a step of reconditioning of the particles before calculating the likelihood function. At this step, the following considerations are done: (1) Particles sufficiently close to the feasible region can be modified in order to belong to the feasible region if such changes interfere only in states related deterministically with the output; (2) The states \( x_1^i \) that have been modified must not interfere in the likelihood calculus of the particle \( X^n \) since this would involve accounts with the impulse function \( \delta(\tau) \) which would lead to infinite likelihoods or even indeterminacies;

Thereby, after the step of reconditioning, the particles \( x_1^i \) and \( x_2^i \) in Fig. 2 become \([3.00 1.20]^T\) and \([3.00 0.30]^T\) respectively, and its likelihoods are both 1.

Thus, the algorithm for particles filtering of max-plus systems can be synthesized as follows:

1. \( k = 0 \);  
2. Initialize \( N \) particles, \( X^0_i, i = 1, \ldots, N \) and its respective weights;  
3. \( k \leftarrow k + 1 \);  
4. Augment each existing particle \( X_{k-1}^i \) to \( X_k^i \) (see Eq. (1a));  
5. Read the measurement \( z_k \);  
6. Recondition;  
7. Update the weights of \( X_k^i \) (see Eq. (7)).
8. If necessary, resample;
9. Estimate $x_k$ (see Eq. (9));
10. Go to item 3;

4. Results
Consider the autonomous system, proposed in [5], described by the following matrices:

$$A = \begin{bmatrix}
\varepsilon & \varepsilon & 4 & \varepsilon & \varepsilon & 2 & \varepsilon & \varepsilon \\
\mu(1,7) & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & \varepsilon \\
\varepsilon & 5 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 1 \\
4 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & \varepsilon \\
\varepsilon & \mu(3,5) & \varepsilon & \mu(1,3) & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\
\varepsilon & \varepsilon & 5 & \varepsilon & 4 & \varepsilon & \varepsilon & \varepsilon \\
\varepsilon & \varepsilon & \varepsilon & 4 & \varepsilon & \varepsilon & \varepsilon & 3 \\
\varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & 5 & \varepsilon & \varepsilon \\
\varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 2 & \varepsilon \\
\varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\
\varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon
\end{bmatrix}$$

$$C = \begin{bmatrix}
\varepsilon & \varepsilon & e & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\
\varepsilon & \varepsilon & \varepsilon & e & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\
\varepsilon & \varepsilon & \varepsilon & \varepsilon & e & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\
\varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & e & \varepsilon & \varepsilon & \varepsilon \\
\varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & e & \varepsilon & \varepsilon \\
\varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & e & \varepsilon \\
\varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & e
\end{bmatrix}$$

It was applied the particle filter in this system and the results were compared to the results of the Observer [4], which determines a lower bound for each state variable. In Fig. 3 can be found the estimated firing sequences for states $x_2$ and $x_5$ produced by the particle filter and by the observer as well as the real firing sequences for these states. In Tab. 1 the estimation errors of the Particle Filter (P.F.) and of the Observer (Obs.) are compared.

The presented results shows that the estimations produced by the particle filter are, in general, closer to the real state than its lower bound calculated by the Observer.

5. Conclusions
We have developed a particle filter algorithm for max-plus systems composed of three key steps which are: (1) the sequential importance sampling, (2) resampling and (3) reconditioning. The choice of importance density adopted in this paper yielded simplicity to the algorithm since the set of particles for a given value of $k$ can be obtained using the equation of dynamic state (see Eq. (1a)) and the set of particles existing in $k - 1$. Although this is a suboptimal choice for the importance density, the results presented shows that the particle filter developed produces estimates close to the real values of the system state. The step of reconditioning the particles allowed us to estimate states associated deterministically with the output, moreover, this step can cause small changes in the density of importance. We are seeking a reconditioning algorithm that interferes even less in the importance density.

References