

Temporal Object Tracking in Large-Scale Production Facilities using Bayesian Estimation^{*}

Karl-Philipp Kortmann^{*} Johannes Zumsande^{*}
Mark Wielitzka^{*} Tobias Ortmaier^{*}

^{*} *Institute of Mechatronic Systems, Leibniz University Hannover, GER*
(e-mail: {kortmann, zumsande, wielitzka,
ortmaier}@imes.uni-hannover.de)

Abstract: Moving towards comprehensive digitalization of production facilities, it is critical to know the location of work pieces, charges, or other objects of interest that change location over time during production. For the case of a limited traceability of these objects, we first present a theoretical approach that performs a recursive Bayesian estimation of the object's location over time based on typical passage measurements in production (e.g. light barriers or RFID systems). The probabilistic method is based on a directed acyclic graph modeling the transfer and sojourn of the objects in the production network. Subsequently, the method is validated on simulated data while varying both size and measurement conditions of the process. The results show the benefit of the proposed method against a single estimation and demonstrate its potential for the application in real time scenarios.

Keywords: production systems, recursive estimation, process models, directed graphs, stochastic systems, stochastic modeling, probabilistic models, Bayesian filter

1. INTRODUCTION

In modern industrial production processes, the tracking of individual work pieces or intermediate products over time is essential to optimize process flows and monitor process quality and stability. For storage and processing of the tracking information, a data warehouse can be set up, which enables the usage of advanced methods of data mining (see Tao et al. (2018)). These methods imply process monitoring (as recently shown by Maurer et al. (2017)), regression and forecasting of relevant information, like energy consumption by Shrouf and Miragliotta (2015), or the data-based control of either single sub-processes or high-level process control as proposed by Răileanu et al. (2018).

Common solutions for the purpose of object recognition and tracking use both hardware sensors (like QR-Codes or wireless RFID-Chips, see Liewald et al. (2018) and Rahmati et al. (2007) respectively) and the mostly deterministic plant control logic to match the information about location and time with the individual tracked object of interest.

In the case that this combined information is not fully available (due to e.g. economical or technical reasons), there is the option of estimating the unobserved events of transition, where transition in this case means the

entrance, exit, or passage of an object with respect to a certain location or state in the process.

In this paper, we present a novel approach of estimating the transition time of production units like work pieces, charges of fluids, or bulk material from one production state to another by a recursive Bayesian estimation under consideration of a process model. As a result, we receive probability distributions of the sojourn of each unit for every discrete location over time.

First, we introduce the chosen mathematical representation of an arbitrary production process in section 2. Following this, a recursive Bayesian update derived from propagated measurements is applied to the problem in section 3. The validation of the method itself and analysis of the process parameter's impact will be presented in section 4.

2. STOCHASTIC PROCESS MODELING

Directed graphs have established themselves as a suitable mathematical structure for describing temporal or logical processes of arbitrary structure. Petri nets are a well-known application of directed graphs and are used in manufacturing, e.g. by Qiao et al. (2015) to analyze the cycle-time in a wafer production process. But also in the area of modeling with uncertainties, stochastic extensions of Petri nets are applied, e.g. by Ammour et al. (2016) for fault prediction in discrete systems.

In automaton theory there are also approaches on probabilistic methods using directed acyclic graphs (DAG), as Niggemann et al. (2012) present a learning algorithm to identify a probabilistic automaton from the control signals

^{*} This work has been financially supported by the European Regional Development Fund (ERDF) grant ZW 6-85018381. We would like to thank our industrial partners Heinrich Meier Eisengießerei GmbH & Co. KG, IAV GmbH and KÜNKELE WAGNER Germany GmbH for their support.

of a mechatronic process. Whereas the latter publication focuses on the detection of anomalies from a process perspective, in this paper the tracking of objects in a production process over time is pursued.

2.1 Stochastic Sojourn Time Graph

The DAG's structure regarding the probabilistic modeling of pathway decision as well as the sojourn time by Niggemann et al. (2012) has recently been adopted for the temporal simulation of the work piece flow in production facilities, see Zumsande et al. (2019).

Similar to the latter, the graph's *vertices* in this paper represent a certain location with temporary stay like transportation, machining, storage or manual manipulation of the traced object (in the following called *token*). The outgoing directed *edges* of a vertex describe all possible transitions a token may take from its current vertex and determine also the conditions like the sojourn time and transition probability. In the following, $t \in \mathbb{R}$ is used in the domain of the real process time while $\tau \in \mathbb{R}_0^+$ to denote the inherently positive sojourn time. Based on this, we define a *stochastic sojourn time graph* (SSTG) as follows:

Definition (stochastic sojourn time graph). A SSTG is a tuple $S = (G, T, s, p, \Delta, c)$, where:

- $G = (V, E)$ is a connected directed acyclic graph. V is an finite set of vertices and $E \subseteq V \times V$ is a set of ordered pairs (i, j) , called edges of vertices $i, j \in V$, where $i \neq j$ (i. e. G is *simple*). G may contain multiple sources (vertices without predecessors)

$$V_{\text{so}} = \{v \in V \mid N^-(v) = \emptyset\} \quad (1)$$

and sinks (vertices without successors)

$$V_{\text{si}} = \{v \in V \mid N^+(v) = \emptyset\} \quad (2)$$

with the set of predecessors $N^-(v)$ and the set of successors $N^+(v)$ of v .

- T is a set of tokens and $s : V_{\text{so}} \rightarrow [0, 1]$ maps a probability to each source, specifying the initial distribution of all tokens, where $\sum_{i \in V_{\text{so}}} s_i = 1$.
- The function $p : E \rightarrow [0, 1]$ maps a probability to each edge, where $\sum_{j \in N^+(i)} p_{i,j} = 1, \forall i \in V \setminus V_{\text{si}}$. ${}_p X_i$ denotes the related discrete random variable of the token's next transition, starting from vertex i .
- The function $\Delta : E \rightarrow \mathcal{P}(\mathbb{R}_0^+)$ assigns a probability distribution from the set of all distributions supported on $[0, \infty)$ to each edge. $\Delta X_{i|j}$ denotes the corresponding random variable (distributed with probability density function (PDF) $\delta_{i|j}(\tau)$) of the sojourn time at Vertex (i) under the condition of a subsequent transition to j . Contrary to this, ΔX_i is the random variable of sojourn time without knowledge of the next transition.
- The capacity $c : V \rightarrow \mathbb{N}$ describes the maximum number of tokens located at a certain vertex at the same time.

As every model, the chosen representation is a trade-off between a preferably high coverage of the most important restrictions and causalities of a real production facility on the one hand and a favorable mathematical abstraction for the later described graph based algorithms on the other

hand.

The capacity of a vertex and the predetermined possible transitions in the process (e.g. serial and parallel production) are examples for these realistic restrictions. By defining the graph as finite and acyclic, path searching can be solved in at least polynomial time (see Chatterjee et al. (2015)). Also a random work piece loss can be modeled by a sink as shown in Fig. 1 at vertex 7.

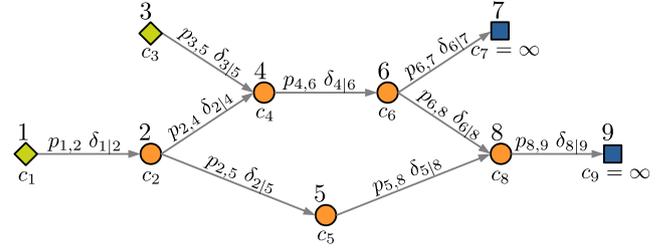


Fig. 1. Example of a simple SSTG with nine vertices and edges each. Sources are marked as green diamonds and sinks as blue squares.

Additionally to the definition of the SSTG, we suppose all capacities c to be infinite to ensure both random variables ΔX_i and ${}_p X_i$ being stochastically independent from the number of tokens located at the set of successive vertices $N^+(i)$. This implies, that tokens are able to overtake each other and do not influence among themselves in their temporal behavior, which is generally not the case for work pieces in serial production processes, but is considered to be a permissible simplification here.

In terms of probability theory, a single transition of one token starting from vertex i can be interpreted as a two-stage random experiment and therefore described by a mixed joint probability distribution (see Ross et al. (1996)) with density

$$f_{{}_p X_i, \Delta X_i}(p x_i, \delta x_i) = P({}_p X_i = p x_i) \cdot f_{\Delta X_i | {}_p X_i}(\delta x_i | p x_i), \quad (3)$$

with the conditional probability density function $f_{\Delta X_i | {}_p X_i}$ and the realization ${}_p x_i$ of ${}_p X_i$ and δx_i of ΔX_i respectively. In terms of the proposed SSTG, (3) can be formulated as

$$f_i(t, j) = p_{i,j} \delta_{i|j}(\tau). \quad (4)$$

Consequently, ΔX_i and ${}_p X_i$ are independent if and only if

$$\delta_{i|j} = \delta_{i|k}, \quad \forall j, k \in N^+(i), k \neq j \quad (5)$$

and dependent otherwise.

2.2 Model Parameters

For the proposed SSTG a number of parameters has to be determined (see table 1). Both the graph's topology and the vertices' capacity are considered to be definite and manually identifiable. Since the edge probabilities $p_{i,j}$ of all outgoing edges of a vertex $i \in V \setminus V_{\text{si}}$ sum to one, all single edges $\{(i, j) \mid |N^+(i)| = 1\}$ hold probability 1 and it is evident that

$$\text{DOF}_p = |E| - |V| + |V_{\text{si}}|, \quad (6)$$

where DOF_p is the degree of freedom regarding the edge probabilities.

In case of the abstracted production process shown in Fig. 1, one edge probability must be determined after vertex 2 and vertex 6 respectively.

Table 1. SSTG model parameters

Symbol	Explanation
$s_v \in [0, 1]$	A Tokens Probability to start from source $v \in V_{so}$
$p_{i,j} \in [0, 1]$	Edge probability
c_i	Capacity of vertex i
$\delta_{i j}(\tau \bullet)$	PDF of the sojourn time from vertex i to j , parameterized by \bullet
$\mu_{i j}$	First moment (mean) of $\delta_{i j}$
$\sigma_{i j}^2$	Second central moment (variance) of $\delta_{i j}$

As described in section 2.1, the edge-specific sojourn time is distributed with density $\delta(\tau)$ and support on \mathbb{R}_0^+ because a negative sojourn time is unfeasible. In literature, the gamma distribution with density $\gamma(\tau)$ is frequently used to model waiting times (see e.g. Sun and Jusko (1998)) and so it is in this paper for later simulation. With its shape parameter α and rate parameter β , it follows

$$\delta_{i|j}^{\text{sim}}(\tau | \alpha_{i|j}, \beta_{i|j}) = \gamma(\tau | \alpha_{i|j}, \beta_{i|j}) \quad (7)$$

$$= \frac{\beta^{\alpha_{i|j}} \tau^{\alpha_{i|j}-1} e^{-\beta_{i|j}\tau}}{\Gamma(\alpha_{i|j})} \quad (8)$$

for $t > 0$, $\alpha_{i|j}, \beta_{i|j} > 0$ and with the gamma function $\Gamma(\cdot)$. Both parameters can be substituted by the more interpretable mean $\mu_{i|j}$ variance $\sigma_{i|j}^2$ of the distribution

$$\alpha_{i|j} = \frac{\mu_{i|j}^2}{\sigma_{i|j}^2}, \quad \beta_{i|j} = \frac{\mu_{i|j}}{\sigma_{i|j}^2}, \quad (9)$$

with $\mu_{i|j} \geq 0$ and $\sigma_{i|j}^2 > 0$.

There are several approaches for the identification of the listed parameters in production processes, in particular see Maier et al. (2011) for the identification of timed hybrid automata and Coit and Jin (2000) for the maximum likelihood estimation of gamma distributions in the course of modeling failure times. Since the parameters for the presented verification were given for the simulation, the identification process is not part of this publication.

2.3 Measurement Propagation

For the later recursive Bayesian estimation, the true state x^ξ is defined as the point in time of the transition of a token h via edge (i, j) (in the time domain of t), with index set

$$\xi = \{(h, (i, j)) | h \in T(i, j) \in E\} \quad (10)$$

Furthermore we define the k -th measurement z_k^ξ as noisy measurement of x^ξ . It is assumed that each measurement can be assigned to one token uniquely. This fits the earlier mentioned application scenarios with hardware like RFID or QR-Readers. We further assume a static white Gaussian measurement noise

$$R_{\text{meas}} \sim \mathcal{N}(0, \sigma_{\text{meas}}^2) \quad (11)$$

with a constant measurement noise variance σ_{meas}^2 and the normal distribution $\mathcal{N}(\cdot)$. Considering the previous, the measurement likelihood function is given by:

$$p(z_k^\xi | x^\xi) = \mathcal{N}\left(z_k^\xi | x^\xi, \sigma_{\text{meas}}^2\right) \quad (12)$$

Taking into consideration that the information provided by a single measurement of one token crossing a specific edge does not only effect our knowledge about the

corresponding true state but also about the state of the remaining edges, we propose the propagation of the single measurement over all edges. With the given parameterized SSTG model, describing the transitional behavior of each token over time t in the process, let $\Phi X_{i,j}^h$ be the random variable of the transition time (e.g. given by a single measurement z_k^ξ) of token h regarding edge (i, j) distributed with density $\phi_{i,j}^h(t)$ and $\Delta X_{j,m}$ be the random variable of the sojourn time of a successive edge $m \in N^+(j)$ with density $\delta_{j,m}(\tau)$. Then the transition time can be propagated towards this successive edge by adding both random variables (cf. Ross et al. (1996))

$$\Phi X_{j,m}^h = \Phi X_{i,j}^h + \Delta X_{j,m} \quad (13)$$

and it follows for the corresponding PDFs

$$\phi_{j,m}^h(t) = (\phi_{i,j}^h * \delta_{j,m})(t), \quad (14)$$

with the convolution operator $*$.

As the PDF of the sum of multiple gamma distributed random variables over a path on G cannot easily be expressed in closed form according to Ansari et al. (2017), in this paper all sojourn times in the propagation step will be approximated by normal distributions. In order to estimate the parameters for the normal distributed sojourn time for the measurement propagation, the *method of moments* is applied and the approximated PDF of the sojourn time results to:

$$\tilde{\delta}_{i,j}^{\text{prop}}(\tau | \mu_{i,j}, \sigma_{i,j}^2) = \mathcal{N}(\tau | \mu_{i,j}, \sigma_{i,j}^2). \quad (15)$$

It should be noted that this computationally necessary approximation has only limited validity in case of a small mean with simultaneously high variance because the normal distribution provides an unfeasible probability density for $t < 0$. Since this work always assumes a sufficiently high mean of the sojourn time in relation to its variance, this is not considered.

In order to propagate a single measurement z_k^ξ recursively, the graph G is partitioned into a forward subgraph $G_j^+ \subseteq G$, including all reachable vertices and edges from vertex j , and a backward subgraph $G_j^- \subseteq G$, including all vertices and edges which could have been crossed before causally. Since G is provided to be acyclic, it directly applies

$$G_j^+ \cap G_j^- = \emptyset, \quad (16)$$

so that no vertex or edge can be part of both subgraphs. In Fig. 2, the exemplary SSTG from section 2.1 is partitioned for an observation at edge $(4, 6)$. As vertex 5 is not reachable anymore, it is not part of either subgraph. The resulting scaled probability density function and cumulative distribution function (CDF) of the propagated measurement over time is shown in Fig. 3, while obtaining $z_{4,6} = 5\text{ s}$ and $\sigma_{\text{meas}}^2 = 1 \times 10^{-3} \text{ s}^2$. The used edge parameters can be found in appendix table A.1.

When a single measurement z_k^ξ is propagated from multiple edges into one vertex (see e.g. vertex 8 in Fig. 1), this merging inevitably results in a PDF consisting a (weighted) sum of normal distributions. Therefore, the likelihood of the propagated measurement (denoted by y_k^ξ to avoid the confusion with the original single measurement z_k^ξ) appears in form of a Gaussian mixture distribution (GMD)

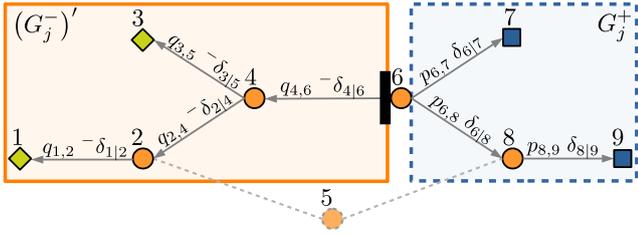


Fig. 2. Partitioning of the example SSTG for a transition measurement across edge (4, 6) (solid black bar). The right square (dashed blue) encloses the subgraph relevant for the forward propagation of the measurement. The left square (solid orange) encloses the reverse subgraph for the backward propagation.

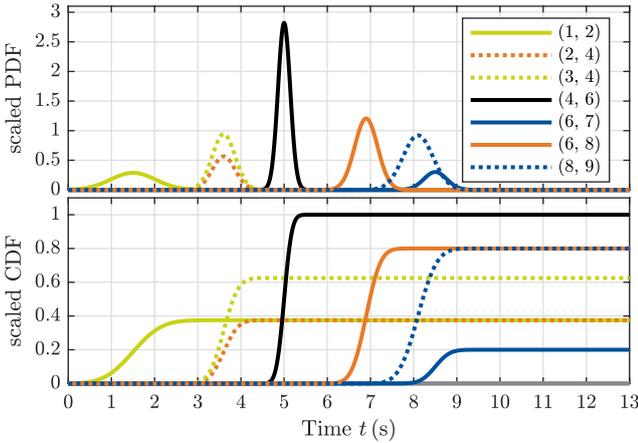


Fig. 3. Scaled PDF (top) and CDF (bottom) of the propagated measurement at edge (4, 6) (black curve) over all reachable edges. (Both PDF and CDF are scaled with the token’s probability of having crossed the specific edge within $t \rightarrow \infty$, considering the appearance of the measurement.)

$$p(y_k^\xi | x^\xi) = \sum_{n=1}^{N_k^\xi} \eta_{k,n}^\xi \mathcal{N}(y_k^\xi | \mu_{k,n}^\xi, \sigma_{k,n}^{2\xi}) \quad (17)$$

$$\text{and } \sum_{n=1}^{N_k^\xi} \eta_{k,n}^\xi = 1,$$

with the likelihood function of the propagated measurement $p(y_k^\xi | x^\xi)$, the number of mixture components N_k^ξ , and the component’s weighting coefficient $\eta_{k,n}^\xi$, mean $\mu_{k,n}^\xi$, and variance $\sigma_{k,n}^{2\xi}$.

3. BAYESIAN ESTIMATION WITH GAUSSIAN MIXTURES

Since the true state x^ξ in this paper is not obtained to be dynamic, we use a recursive Bayesian estimation of the static transition time by a sequence of propagated measurements. Therefore, the prediction step from Bayesian filtering (known in particular for its most frequent use in Kalman filters) is dropped, as there is no model function or process noise to be included. According to Särkkä (2013), the solution to the estimation problem (also called *param-*

eter estimation, as the estimated state is considered to be static) is formulated as follows:

- (1) The propagated measurements y_k^ξ are obtained to be conditionally independent and modeled by their likelihood function $p(y_k^\xi | x^\xi)$, given by (17).
- (2) All information in the beginning of the estimation is given by the initial prior

$$p(x^\xi) = \mathcal{N}(x^\xi | \mu_0^\xi, \sigma_0^{2\xi}), \quad (18)$$

with the initial mean μ_0^ξ and initial variance $\sigma_0^{2\xi}$.

- (3) The estimation is performed recursively after each k -th measurement of token h with an update of the posterior distribution

$$p(x^\xi | z_{1:k}^\xi) = \frac{p(z_k^\xi | x^\xi) p(x^\xi | z_{1:k-1}^\xi)}{\int p(z_k^\xi | x_k^\xi) p(x^\xi | z_{1:k}^\xi) dx^\xi}, \quad (19)$$

where the denominator is often called *model evidence* and serves as a scalar normalization term.

3.1 Application and Reduction of Gaussian Mixtures in Bayesian Estimation

In this paper, the frequently used closed form solution for the recursive Bayesian estimation by enforcing the posterior distribution to be a Gaussian mixture distribution (GMD) is applied as we obtain the likelihood of our propagated measurements to be a GMD (17). This method was first introduced by Sorenson and Alspach (1971) and has been extended frequently, most notably to the presented problem by Vo and Ma (2006). Like the authors, we tackle the problem of a time varying number of observed targets, since single states can vanish as Fig.3 demonstrated for edges (2, 5) and (5, 8).

With the determination of the prior (18) being plainly normal distributed, every updated posterior distribution is also a GMD and all of its weights, means and variances can be calculated in closed form, see Vo and Ma (2006).

As Sorenson and Alspach (1971) pointed out early, the number of Gaussian components in the a posteriori distribution grows exponentially with k and the recursive Bayesian update, given by (19), has therefore a time complexity of $\mathcal{O}(N^k)$ with the presumed maximum number of GMD components N of the measurement likelihood (17) within k steps.

Since then, a number of reduction algorithms has been proposed and compared (see Crouse et al. (2011)), where in this paper the method by Runnalls (2007) is used, which minimizes an upper bound regarding the Kullback-Leibler divergence between the full and the reduced mixture.

In terms of the following validation, $\hat{x}_{i,j}^h(t) = p(x^\xi | z_{1:\kappa}^\xi)$ denotes the final PDF of the posterior for edge (i, j) (i. e. after the maximum number of measurements κ regarding token h). The mean of this PDF is indicated by $\hat{x}_{i,j}^h$, which is used as point estimate for the transition time in the next section.

4. VALIDATION AND PARAMETER ANALYSIS

For the validation of the proposed methods the flow of tokens will be simulated within three scenarios with different complexity of the graph's topology, determined by the number of vertices and edges in each of the three cases. As preliminary experiments showed a strong dependency between the method's performance and the selected topology and parameterization of the graph, 10 independent random DAGs per scenario are drawn under given specification of $|V|$, $|E|$, $|V_{\text{so}}|$ and $|V_{\text{si}}|$, see table 2.

The SSTG model parameters, shown in table 1, are randomly drawn with determining $\mu X_{i,j} \sim \gamma(10, 10)$ and $\sigma^2 X_{i,j} \sim \gamma(10, 10)$, where the random variable $\mu X_{i,j}$ represents the sojourn time distribution's mean of edge (i, j) and $\sigma^2 X_{i,j}$ its variance. This is done to ensure $E[\mu X_{i,j}] = E[\sigma^2 X_{i,j}] = 1$ s and analogically $\text{Var}(\mu X_{i,j}) = \text{Var}(\sigma^2 X_{i,j}) = 0.1$ s² for better comparability of the validation metrics. The capacity of all vertices is assumed to be infinite as described in section 2.1 and the measurement noise's variance is determined to be $\sigma_{\text{meas}}^2 = 0.01$ s². Each prior PDF is distributed with

$$p(x_1^n) \sim \mathcal{N}(10 \text{ s}, 100 \text{ s}^2). \quad (20)$$

For each randomly drawn graph the flow of 1000 tokens is simulated while calculating the individual sojourn times by (7). The number of components allowed within the GMD of an updated posterior distribution is limited to 20 by Runnalls' method (cf. section 3.1).

Table 2. Specification of the topology for the validation scenarios

	Scenario A	Scenario B	Scenario C
$ V $:	10	20	40
$ E $:	20	40	80
$ V_{\text{so}} $:	2	3	4
$ V_{\text{si}} $:	2	3	4

4.1 Variation of Measurement Parameters

Besides the influence of topology the effect of both parameter noise and the number of measurement stations is analyzed.

Since the proposed method relies on the knowledge of the SSTG model parameters, these have to be identified for real production facilities. In order to analyze the influence of an uncertainty regarding the model parameters, four levels of white Gaussian noise added to each parameter group and quantified by a decreasing signal-to-noise ratio from $\text{SNR}_{\infty \text{ dB}}$ to $\text{SNR}_{10 \text{ dB}}$ are investigated.

By varying the number of edges that measure the transition of a token, the realistic production case of an only partially observed process (e.g. by a few single RFID readers) will be reflected. A distinction is made between a randomly chosen set of edges with measurement capability (denoted by $E_{\text{meas}} \subseteq E$) of either 10%, 20%, or 50% of $|E|$. In addition, it is considered that only one transition per token is measured (more precisely: Its last transition into a sink). This case should serve as a benchmark, since there is exactly one Bayesian update performed on the initial prior regarding the transition statistics given by the SSTG.

4.2 Evaluation Metrics

All three scenarios with varying model parameter noise and number of measurement stations are validated regarding three metrics. The average error of the method's point estimation and its distribution is quantified by the root-mean-square error

$$\text{RMSE} = \sqrt{e_{i,j}^h{}^2} \quad (21)$$

with point estimation error $e_{i,j}^h = \hat{y}_{i,j}^h - y_{i,j}^h$ and an application-specific mean function (denoted by an overline)

$$\bar{\cdot} = \frac{1}{|T|} \sum_{h \in T} \left(\frac{1}{|E^h|} \sum_{(i,j) \in E^h} \cdot \right), \quad (22)$$

with the set of edges $E^h \subseteq E$ transitioned by token h during the simulation.

To take not only a point estimation but also the whole distribution of the estimation into account, we use the Bhattacharyya distance

$$D_B(f, g) = -\ln \left(\int_{-\infty}^{\infty} \sqrt{f(t)g(t)} \right) \quad (23)$$

as measure of the similarity (cf. Bi et al. (2017)) between the estimated PDF $f(t) = \hat{y}_{i,j}^h(t)$ and the simulated transition time $g(t) = \mathcal{N}(t | y_{i,j}^h, \sigma_{\text{meas}}^2)$. To give $g(t)$ support on \mathbb{R} , it is assumed to be normal distributed with equal measurement noise σ_{meas}^2 .

In order to evaluate not only the goodness of the estimated transition time but also the accuracy of a token's path which is estimated to be the most likely one, we define the paths accuracy

$$\text{ACC} = \frac{1}{|T|} \sum_{h \in T} \mathbf{1}(\mathcal{P}_h = \hat{\mathcal{P}}_h), \quad (24)$$

where \mathcal{P}_h is the sequence of edges the token h visited while simulation and $\hat{\mathcal{P}}_h$ is the most probable sequence of edges after the last update of the Bayesian estimation. $\mathbf{1}(\cdot)$ denotes the indicator function.

Finally, the average computing time for an update step (including measurement propagation and eventual GMD reduction) is given by t_{upd} . The validation was performed on a single core CPU with 3.5 GHz and 16 GB RAM in the MATLAB R2019A computing environment.

4.3 Results

The results shown in table 3 reflect a clear trend regarding the three scenarios. For all three quality metrics, the value deteriorates as the complexity of the graph increases (that means a rising value in case of RMSE and mean Bhattacharyya distance \bar{D}_B and a decreasing value for the accuracy). In particular, the calculation time increases even more than linear with the number of edges in the graph. This was expected, since the computing time without reduction mechanism of the GMD even grows exponentially with the number of measurements (cf. section 3.1).

Concerning the number of measuring edges, the validation result indicates a consistent decrease for both RMSE and Bhattacharyya distance, in case of an increasing amount of

Table 3. Validation results with simulated token flow for three different graph sizes (scenarios). The estimation was performed with a varying number of measuring stations (edges) and different noise levels in the model parameters.

		RMSE (s)			$\overline{D_B}$ (-)			ACC (-)			t_{upd} (ms)		
		A	B	C	A	B	C	A	B	C	A	B	C
SNR _∞ dB	$E_{\text{meas},1}$	4.23	7.88	13.49	0.61	0.95	2.39	59.2	43.8	22.1	84	198	691
	$E_{\text{meas},10\%}$	3.99	7.48	11.81	0.53	0.84	2.10	73.5	54.3	25.2	113	250	781
	$E_{\text{meas},20\%}$	3.50	6.71	11.19	0.52	0.82	2.06	80.2	59.3	29.5	149	399	1280
	$E_{\text{meas},50\%}$	2.99	5.81	9.40	0.44	0.64	1.75	99.3	93.2	44.4	190	557	1562
SNR ₄₀ dB	$E_{\text{meas},1}$	4.54	8.38	14.01	0.64	1.00	2.51	60.3	45.1	21.5	95	198	748
	$E_{\text{meas},10\%}$	4.10	7.22	13.40	0.58	0.88	2.24	71.0	55.1	26.3	89	240	729
	$E_{\text{meas},20\%}$	4.00	6.83	12.53	0.52	0.86	2.23	72.2	58.5	28.3	138	380	1386
	$E_{\text{meas},50\%}$	3.20	5.85	9.66	0.44	0.72	1.76	99.3	80.9	40.1	193	469	1965
SNR ₂₀ dB	$E_{\text{meas},1}$	4.83	8.87	15.62	0.65	1.05	2.64	54.0	40.7	19.6	74	221	645
	$E_{\text{meas},10\%}$	4.24	7.95	13.96	0.60	0.91	2.48	60.4	47.5	23.6	87	271	835
	$E_{\text{meas},20\%}$	4.17	8.00	12.58	0.58	0.88	2.30	68.1	54.6	25.2	147	390	1104
	$E_{\text{meas},50\%}$	3.47	6.13	10.58	0.48	0.74	1.91	99.3	88.6	40.6	222	549	2040
SNR ₁₀ dB	$E_{\text{meas},1}$	5.64	10.73	19.25	0.79	1.27	3.15	48.6	33.6	17.0	78	184	630
	$E_{\text{meas},10\%}$	5.21	9.97	17.94	0.74	1.18	2.92	51.0	40.6	22.6	112	229	925
	$E_{\text{meas},20\%}$	5.39	9.55	17.01	0.72	1.05	2.61	68.4	49.6	22.7	169	306	1425
	$E_{\text{meas},50\%}$	4.13	8.13	13.72	0.56	0.87	2.19	94.4	67.7	35.4	225	592	1655

measurements. In terms of the accuracy, the proportion of correctly estimated token paths could be increased by an average of 26 percentage points. Even in case of the highest noise of the model parameters within the validation setup, a clear improvement can be seen in all three metrics.

As it was to be expected, there is a negative influence on the estimation quality for increasing noise in the model parameters. While the effect from noiseless (SNR_∞ dB) to SNR₄₀ dB is hardly noticeable due to the logarithmic scale of the signal-to-noise ratio, in the case of SNR₁₀ dB a significant deterioration is visible for all three metrics. The computing time for one update is uninfluenced by the signal-to-noise ratio, since the imprecision of the parameters affects the estimation quality but not the number of calculations per update step.

In comparison with the trivial case of a single measurement ($E_{\text{meas},1}$), it can be observed, that the estimation quality is increased in all cases, so that the positive benefit of the methodology seems reasonable.

The measured computing time appears tolerable for most real-scenario online or real-time applications but depends essentially on the complexity of the graph as well as the number of measurements. Therefore, it must be investigated in each individual case whether the particular real-time requirements can be met.

5. CONCLUSION AND FUTURE WORK

In this paper, a novel method for combining a graph-based, stochastic model of an arbitrary production process with a recursive Bayesian estimation was elaborated with the aim of estimating the transition times of objects between different process steps, even if there is only partial information about their temporal location available. It was shown, how a single temporal measurement can be propagated due to the proposed stochastic sojourn time graph. The proposed method was validated with simulated data, covering different levels of graph complexity, a varying proportion of measuring edges, and four distinct noise levels of the model parameters.

The results have revealed that the method can improve the transition time estimation with a recursive Bayesian update compared to a static estimation. Also a real-time application seems reasonable due to the measured computing time.

Nevertheless, the examination in a real application case is still pending and indispensable, since in the presented case only simulated data were used and a benefit of the results through the methodological proximity between simulation method and estimation model cannot be excluded. More than one assumption, however, results from simplifications, the validity of which should be validated in reality. In addition, other more sophisticated Bayesian estimation methods, such as the particle filter or an unscented Kalman filter, should also be examined for the presented problem.

REFERENCES

- Ammour, R., Leclercq, E., Sanlaville, E., and Lefebvre, D. (2016). Faults prognosis using partially observed stochastic petri nets. In *2016 13th International Workshop on Discrete Event Systems (WODES)*, 472–477. IEEE.
- Ansari, I.S., Yilmaz, F., Alouini, M.S., and Kucur, O. (2017). New results on the sum of gamma random variates with application to the performance of wireless communication systems over nakagami-m fading channels. *Transactions on Emerging Telecommunications Technologies*, 28(1), e2912.
- Bi, S., Prabhu, S., Cogan, S., and Atamturktur, S. (2017). Uncertainty quantification metrics with varying statistical information in model calibration and validation. *AIAA Journal*, 3570–3583.
- Chatterjee, S., Paladhi, S., Hore, S., and Dey, N. (2015). Counting all possible simple paths using artificial cell division mechanism for directed acyclic graphs. In *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*, 1874–1879. IEEE.
- Coit, D.W. and Jin, T. (2000). Gamma distribution parameter estimation for field reliability data with missing failure times. *Iie Transactions*, 32(12), 1161–1166.

Crouse, D.F., Willett, P., Pattipati, K., and Svensson, L. (2011). A look at gaussian mixture reduction algorithms. In *14th International Conference on Information Fusion*, 1–8. IEEE.

Liewald, M., Karadogan, C., Lindemann, B., Jazdi, N., and Weyrich, M. (2018). On the tracking of individual workpieces in hot forging plants. *CIRP Journal of Manufacturing Science and Technology*, 22, 116–120.

Maier, A., Niggemann, O., Vodencarevic, A., Just, R., and Jaeger, M. (2011). Anomaly detection in production plants using timed automata. In *8th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, 363–369.

Maurer, I., Riva, M., Hansen, C., and Ortmaier, T. (2017). Cloud-based plant and process monitoring based on a modular and scalable data analytics infrastructure. In *Tagungsband des 2. Kongresses Montage Handhabung Industrieroboter*, 33–42. Springer.

Niggemann, O., Stein, B., Vodencarevic, A., Maier, A., and Büning, H. (2012). Learning behavior models for hybrid timed systems. In *26th AAAI Conference on Artificial Intelligence (AAAI-12)*, 1083–1090.

Qiao, Y., Wu, N., Zhu, Q., and Bai, L. (2015). Cycle time analysis of dual-arm cluster tools for wafer fabrication processes with multiple wafer revisiting times. *Computers & Operations Research*, 53, 252–260.

Rahmati, A., Zhong, L., Hiltunen, M., and Jana, R. (2007). Reliability techniques for rfid-based object tracking applications. In *37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN'07)*, 113–118. IEEE.

Răileanu, S., Anton, F., Borangiu, T., Anton, S., and Nicolae, M. (2018). A cloud-based manufacturing control system with data integration from multiple autonomous agents. *Computers in Industry*, 102, 50–61.

Ross, S.M., Kelly, J.J., Sullivan, R.J., Perry, W.J., Mercer, D., Davis, R.M., Washburn, T.D., Sager, E.V., Boyce, J.B., and Bristow, V.L. (1996). *Stochastic processes*, volume 2. Wiley New York.

Runnalls, A.R. (2007). Kullback-leibler approach to gaussian mixture reduction. *IEEE Transactions on Aerospace and Electronic Systems*, 43(3), 989–999.

Särkkä, S. (2013). *Bayesian filtering and smoothing*, volume 3. Cambridge University Press.

Shrouf, F. and Miragliotta, G. (2015). Energy management based on internet of things: practices and framework for adoption in production management. *Journal of Cleaner Production*, 100, 235–246.

Sorenson, H.W. and Alspach, D.L. (1971). Recursive bayesian estimation using gaussian sums. *Automatica*, 7(4), 465–479.

Sun, Y.N. and Jusko, W.J. (1998). Transit compartments versus gamma distribution function to model signal transduction processes in pharmacodynamics. *Journal of pharmaceutical sciences*, 87(6), 732–737.

Tao, F., Qi, Q., Liu, A., and Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems*, 48, 157–169.

Vo, B.N. and Ma, W.K. (2006). The gaussian mixture probability hypothesis density filter. *IEEE Transactions on signal processing*, 54(11), 4091–4104.

Zumsande, J., Kortmann, K.P., Wielitzka, M., and Ortmaier, T. (2019). Probabilistic simulation and deter-

mination of sojourn time distribution in manufacturing processes. In *2019 IEEE International Conference on Mechatronics and Automation (ICMA)*, 2094–2099.

Appendix A. PARAMETERIZATION OF EXEMPLARY SSTG

Table A.1. Edge parameterization of the exemplary SSTG in Fig. 1

(i, j)	$p_{i,j}$	$\mu_{i,j}$	$\sigma_{i,j}^2$
(1, 2)	1.0	4.9	0.1
(2, 4)	0.6	2.1	0.2
(2, 5)	0.4	4.3	0.1
(3, 4)	1.0	3.2	0.1
(4, 6)	1.0	1.4	0.1
(5, 8)	1.0	1.0	0.1
(6, 7)	0.2	3.5	0.1
(6, 8)	0.8	1.9	0.1
(8, 9)	1.0	1.2	0.1