HAL open science

Statistical evaluation of Hidden Markov Models topologies, based on industrial synthetic model

Bernard Roblès, Manuel Avila, Florent Duculty, Pascal Vrignat, Frédéric Kratz

To cite this version:

Bernard Roblès, Manuel Avila, Florent Duculty, Pascal Vrignat, Frédéric Kratz. Statistical evaluation of Hidden Markov Models topologies, based on industrial synthetic model. INCOM 2012, 14th IFAC Symposium on Information Control Problems in Manufacturing Information Control Problems in Manufacturing, May 2012, Bucharest, Romania. pp.1029 - 1034, 10.3182/20120523-3-RO-2023.00052 . hal-00706775

HAL Id: hal-00706775

https://hal.science/hal-00706775v1

Submitted on 11 Jun 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Statistical evaluation of Hidden Markov Models topologies, based on industrial synthetic model

Bernard Roblès * **Manuel Avila** * **Florent Duculty** *
Pascal Vrignat * **Frédéric Kratz** **

* *University of Orleans, PRISME Laboratory IRAUS team, 2 Avenue Mitterrand 36000 Châteauroux, FRANCE (e-mail: Bernard.Robles@univ-orleans.fr, Manuel.Avila@univ-orleans.fr, Florent.Duculty@univ-orleans.fr, Pascal.Vrignat@univ-orleans.fr).*
** *PRISME Laboratory IRAUS team, ENSI 88 boulevard Lahitolle 18020 Bourges Cedex, FRANCE (e-mail: Frederic.Kratz@ensi-bourges.fr).*

Abstract: Prediction of physical particular phenomenon is based on knowledges of the phenomenon. Theses knowledges help us to conceptualize this phenomenon throw different models. Hidden Markov Models (HMM) can be used for modeling complex processes. We use this kind of models as tool for fault diagnosis systems. Nowadays, industrial robots living in stochastic environment need faults detection to prevent any breakdown. In this paper, we wish to evaluate three Hidden Markov Models topologies of Vrignat et al. (2010), based on upstream industrial synthetic Hidden Markov Model. Our synthetic model gives us simulation such as real industrial Computerized Maintenance Management System. Evaluation is made by two statistical tests. Therefore, we evaluate two learning algorithms: Baum-Welch Baum et al. (1970) and segmental K-means Viterbi (1967). We also evaluate two different distributions for stochastic generation of synthetic HMM labels. After a brief introduction on Hidden Markov Model, we present some statistical tests used in current literature for model selection. Afterwards, we support our study by an example of simulated industrial process by using synthetic HMM. This paper examines stochastic parameters of the various tested models on this process, for finally come up with the most relevant model and the best learning algorithm for our predictive maintenance system.

Keywords: Hidden Markov Models, model selection, learning algorithms, statistical test, predictive maintenance, Aspin-Welch test, Kolmogorov-Smirnov test, p-value.

1. INTRODUCTION

Fault diagnostics techniques can reduce maintenance downtime. According to Vrignat et al. (2010), we can find two keywords in maintenance definition: *maintain* and *restore*. The first one refers to preventive action. The second refers to corrective action. Thus, maintenance optimization for reliability determines "optimal" preventive maintenance. Events preceding a problem in maintenance activities are often recurrent. Special events series should inform us on next failure. For example, in mechanical systems, noises, vibrations precede failure. The loss of performances reflects failure or technical faults. Our works, Vrignat et al. (2010), have shown that it was possible to model degradation levels of a process and results have shown that our approach combined with work of Zille et al. (2007) can provide decision support for industrial maintenance. We also showed in Vrignat et al. (2010) that our model provided good failure prediction. Moreover, we make a reference model, named *synthetic model*, which fits to real industrial processes. Our goals consist here to evaluate three different Hidden Markov Models topologies, with parameters outcome from this industrial *synthetic model*. The paper deals with the measurement of relevance on Hidden Markov Models parameters, based on several statistical criteria used in current literature.

The objective of this work is to evaluate the best Hidden Markov Model, by using Kolmogorov-Smirnov test (Wang and Schaalje (2009), Drezner et al. (2010)) and Aspin-Welch test (Welch (1951), Held et al. (2010)). Calculation is made with states and observations of three different HMM topologies (figure 2). Symbols (or observations) are produced by a synthetic HMM (the reference model), using two different learning algorithms and two different distributions of symbols.

The structure of this paper is as follows: in section 2, we outline hidden Markov model and define its parameters. We present criteria used to evaluate relevance of HMM parameters, in section 3. We show the evaluation process in section 4. Finally, we use our synthetic model to compare several HMM topologies, from among a candidate set, with previous criterion and then we try to give the best one, in section 5.

2. HIDDEN MARKOV MODEL

We have chosen Hidden Markov Model to describe industrial maintenance events. We first recall what are HMM: A Hidden Markov Model (Rabiner (1989), Fox et al. (2006)) is an automaton with hidden states which consists of unobservable variable. This one represents the system status to be modeled. Only output variable is observable. Moreover, we get observations sequence from output of the automaton. From now, we rename observations sequence as *symbols*, representing these observations (see an example of model topology in figure 1). This is precisely relevance of these symbols that we attempt to evaluate. Hidden Markov Model is characterized by:

- State number;
- Number of distinct observation symbols per state, observation symbols corresponding to the physical output of the system being modeled;
- Distribution probability of state transitions;
- Distribution probability of observation symbols;
- Initial state distribution.

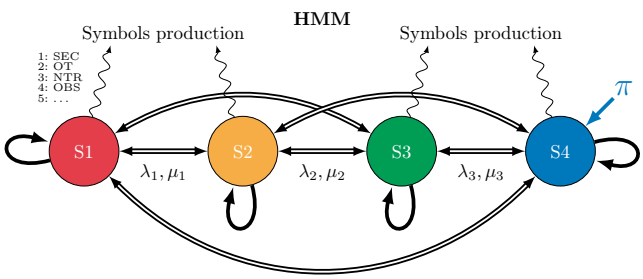


Fig. 1. Four states Hidden Markov Model.

2.1 Markov Assumption

Basically, states prediction is not made more accurate by adding a priori knowledge information, i.e. all useful information for future prediction is contained in present state of the process.

$$P(X_{n+1} = j | X_0, X_1, \ldots, X_n = i) = P(X_{n+1} = j | X_n = i). \tag{1}$$

2.2 Definitions for discrete Hidden Markov Model

We now describe variables for HMM:

- Let N, the number of workable hidden states and $E = \{E_1, E_2, \ldots, E_N\}$, the set of this variable. Let q_t, the value of this variable at time t;
- Modeled process, must match to first-order Markov assumption (see §2.1);
- Let T, the full number of observation symbols and let $X = \{x_1, x_2, \ldots, x_T\}$, observations sequence of the modeled process;
- Let $A = \{a_{ij}\}$, distribution probability of state transitions with:
$$a_{ij} = P(q_{t+1} = E_j | q_t = E_i) \quad 1 \leq i, j \leq N, \tag{2}$$
- Let $B = \{b_j(m)\}$, distribution probability of observation symbol in j state, with:
$$b_j(m) = P(X_t = x_m | q_t = E_j) \quad 1 \leq j \leq N \;\; 1 \leq j \leq T, \tag{3}$$
with X_t, value of observation variable at time t.
- Let $\pi = \{\pi_i\}$, initial state distribution with:
$$\pi = P(q_1 = E_i) \qquad 1 \leq i \leq N, \tag{4}$$
- Hidden Markov Model will be set as: (A, B, π).

3. STATISTICAL TESTS

Most statistical tests assume that samples are taken randomly to achieve Steinebach (2006). This sounds easy but it is actually quite difficult to achieve.

3.1 NIST[1] Tests

In Rukhin et al. (2010), authors propose a statistical package consisting in 15 different tests. These tests were developed to test randomness of random number generators. NIST has verified the performance of these tests using a Kolmogorov-Smirnov test of uniformity on the p-values [2] (see § 3.2). The purpose of this test is to determine whether the number of ones and zeros in a sequence are approximately the same as would be expected for a truly random sequence.
In our study, we use the *frequency test* of the NIST. This test validates that our synthetic model gives real stochastic symbols. The Decision Rule of the test, at the 1% Level, is: If the computed p-values is < 0.01, then we conclude that the sequence is non-random. Otherwise, we conclude that the sequence is random.

3.2 Kolmogorov-Smirnov test

Kolmogorov-Smirnov test is a statistical test that may be used to determine if a set of data comes from a particular probability distribution (Rukhin et al. (2010), Bercu and Chafaï (2007)).

Empirical distribution function $F_n(x)$ for $X_1, \ldots, X_n$ sample is defined by:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i \leq x}, \tag{5}$$

$$\delta_{X_i \leq x} = \begin{cases} 1 \text{ si } & X_i \leq x, \\ 0 & \text{sinon}. \end{cases}$$

The Kolmogorov-Smirnov test statistic is defined as follows:

$$D_n = \sup_x |F_n(x) - F(x)|. \tag{6}$$

[1] National Institute of Standards and Technology
[2] The probability (under the null hypothesis of randomness) that the chosen test statistic will assume values that are equal to or worse than the observed test statistic value when considering the null hypothesis. The p-value is frequently called the "tail probability".

3.3 Aspin-Welch test

Aspin-Welch's test (Welch (1947), Welch (1951)), is defined by t statistic in the following formula:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\sigma^2(\frac{1}{n_1} + \frac{1}{n_2})}}, \tag{7}$$

$$\sigma^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1 + n_2 - 2}, \tag{8}$$

- $\overline{x}_i$: the ith sample mean,
- σ: an estimator of the common standard deviation of the two samples,
- σ_i: samples standard deviation,
- n_i: sample size.

4. EVALUATION PROCESS

We have used synthetic model to produce about 1000 data events. These simulated symbols, according to real industrial process, are obtained by using uniform and normal distribution. Correlatively, we produced states for others topologies by using the same process. Afterwards, these states are used to compare states models. Insofar as states are obtained by different learning and decoding algorithms (diagram of this process is given in figure 3):

- Baum-Welch learning (Baum et al. (1970)), decoding by Forward Variable (Rabiner (1989)),
- Segmental K-means learning (Juang and Rabiner (1990)), decoding by Viterbi, (Viterbi (1967)).

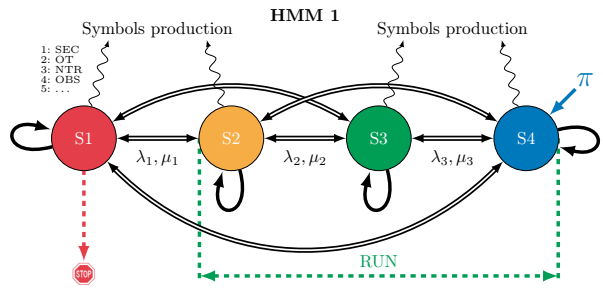
4.1 HMM topologies

Our studied models are represented by automata with four oriented states. This stochastic automata represent the degradation level of an industrial process, S4 to S1, see figure 2. {S4, S3, S2} states, when process is running ("RUN"), and {S1} state, when process is stopped ("STOP"). Topology 1 figure 2(a), describes all possible transitions. With topology 2 figure 2(b), we need to go through all states (S2 and S3) to go from hight level of availability (S4) to low level of availability (S1). Figure 2(c) shows the difference between topologies 2 and 3: S1 becomes a first state of breakdown.

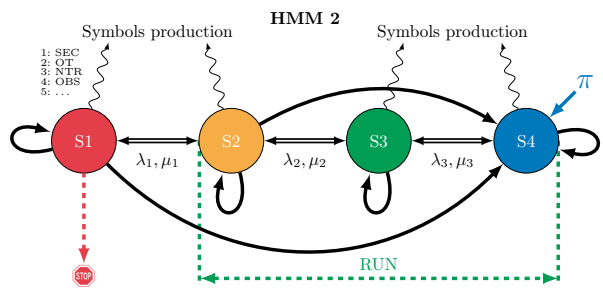
4.2 Simulated industrial CMMS[3]

Nowadays, every industrial factory uses **preventive maintenance**. Maintenance operators consign their actions and observations in a centralized database. We show an example of such database in table 1. For instance, symbols "PM, OT, SP, ..." could characterize maintenance activities carried out on industrial process. We recall the meaning of selected symbols resulting from observations, in table 2. "SP" symbol corresponds to a stop of production units: process state = "STOP" in table 2. It is a critical condition that our research tries to minimize. Process state = "RUN" when production units are running without failure. We study here this kind of **maintenance** by using **synthetic model** (§4.4) to simulate real industrial environment. We have chosen "λ_i" (failure rate) and "μ_i" (repair rate) of HMM parameters (Vrignat (2010)), to match as possible, with maintenance recording (table 1).

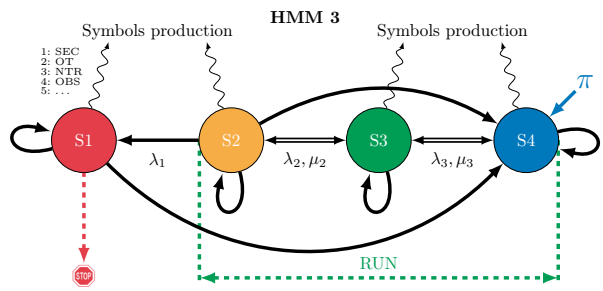
[3] Computerized Maintenance Management System



(a) Topology 1



(b) Topology 2



(c) Topology 3

Fig. 2. Four states Hidden Markov Models, (Vrignat et al. (2010))

Name	Date	Ope.	Cd	IT	N	Code
Dupond	11/01/2007	Lubrication	PM	20	1	9
Dupond	11/01/2007	Lubrication	PM	20	2	9
Dupond	12/01/2007	Lubrication	SEC	30	3	5
Dupond	12/01/2007	Lubrication	PM	30	4	5
Dupond	13/01/2007	Padlock	PM	10	5	6
Dupond	13/01/2007	Padlock	NTR	30	6	5
Dupond	13/01/2007	Padlock	NTR	30	7	5
Dupond	16/01/2007	Lubrication	SP	90	8	1
Dupond	19/01/2007	Padlock	OT	10	9	3

Table 1. Example of recorded events from a maintenance database.

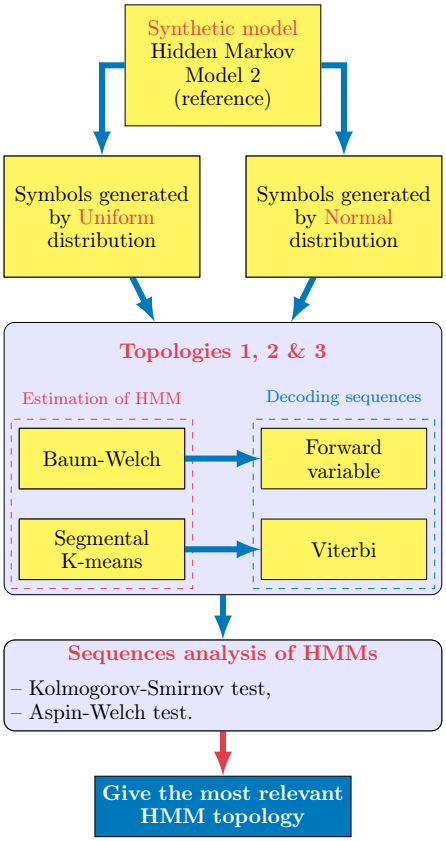


Fig. 3. Matching model method, using synthetic model

Process states	
RUN	
STOP	

	Interventions type	
1	SP	(Troubleshooting / Stop Production)
2	SM	(Setting Machine)
3	OT	(Other)
4	OBS	(Observation)
5	PM	(Preventive Maintenance, Production not stopped)
6	SEC	(Security)
7	PUP	(Planified Upgrading)
8	CM	(Cleaning Machine)
9	PMV	(Preventive Maintenance Visit)
10	NTR	(Nothing to report)

Table 2. Symbolic coding system of maintenance interventions.

4.3 Industrial application

We try to understand the "signature" of industrial CMMS by using HMM topologies. Furthermore, our approach can provide decision support for organizing daily maintenance. Our results presented in Vrignat et al. (2010), provide good estimation of failure probability. Hence, we can help experts to schedule maintenance activities.

4.4 Synthetic model

We have made our synthetic model with Matlab by using four states oriented topology 2 presented in figure 2(b). We have used this model feature because it has good performance in maintenance activities (Vrignat et al. (2010)). Then, we built sequences of data (also named "signature") using this model as the reference model, by injecting stochastic symbols in this HMM. We have used these symbols sequences as Markov chain (see table 3), to model

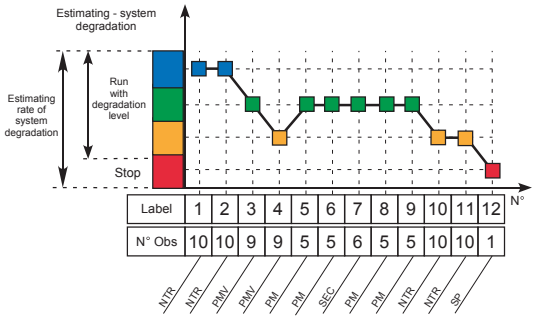


Fig. 4. Degradation of process.

degradation level of a process (example in figure 4). These

PM	PM	SEC	PM	PM	NTR	NTR	SP	...

Table 3. Sequence of a message from maintenance database.

simulated symbols, according to real industrial process (Vrignat et al. (2010)), are obtained by using uniform and Gaussian distribution. We have used these symbols to train three different HMM topologies, described in figure 2, by using two different learning and decoding algorithms: Baum-Welch learning, decoding by Forward Variable, and Segmental K-means learning, decoding by Viterbi as we seen before.

1000 symbols were produced by reference model (distribution in figures 5 and 6). Each sequence ends with a stop of production (symbol SP in red) see figure 4. Then, we got 11 sequences in our 1000 simulated symbols. You can see distribution symbols/states for the first sequence: HMM 1, HMM 1/Baum-Welch and HMM 1/Segmental K-means algorithms, in figure 7. Finally, we obtained states sequences for each HMM outside. Later, these states are used to make comparisons between 3 different HMM topologies (figure 2), with statistical tests studied in section 3. Results are shown in section 5. Diagram of our evaluation process is given in figure 3.

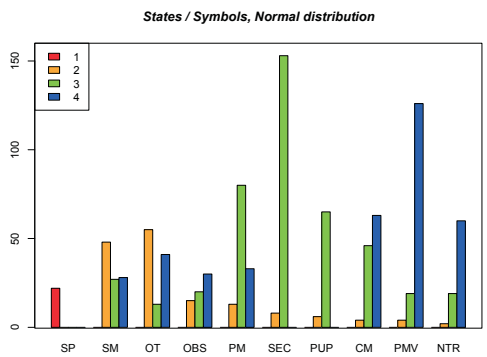


Fig. 5. HMM sequences example, Normal distribution

5. RESULTS

First, we have verified the randomness of stochastic states generated by the synthetic model (figure 5 and 6). The random number generator is tested by *frequency test* § 3.1.

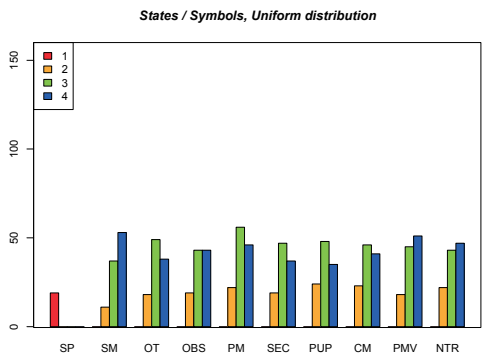


Fig. 6. HMM sequences, Uniform distribution

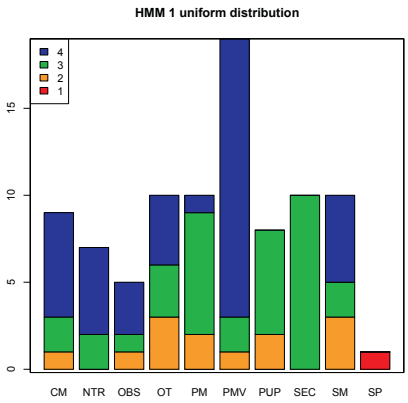
NIST Test	P-value	
HMM	Uniform law	Normal Law
Topology 1	0.47	0.06
Topology 2	0.30	0.02
Topology 3	0.47	0.06

Table 4. p-value of states generated by synthetic model

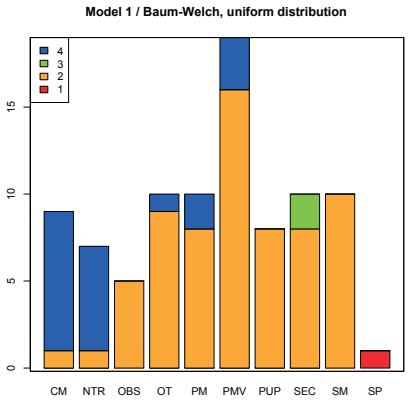
The results in table 4, show that all p-value ≥ 0.01 for all HMM. Then we can consider that sequences of the generator are random enough to apply others statistical tests. Therefore we applied statistical tests on our three Hidden Markov Models topologies figure 2. Aspin-Welch figure 8(a) and Kolmogorov-Smirnov figure 8(b) give the same results: most relevant model is *topology 2*, *Baum-Welch learning algorithm* with Forward Variable decoding is the best learning algorithm and finally, stochastic symbols generated with *normal distribution* is the best way. Kolmogorov-Smirnov test determines if two datasets differ significantly. It has the advantage of making no assumption about data distribution. This test is less sensitive than Aspin-Welch test. This last one is intended for use with samples having unequal variances.
Therefore, without a priori knowledge, we can give the most relevant model in the way of statistical tests. Namely, we verified that the best model, which provides the better estimation of degradation level according to Vrignat et al. (2010), fits to real industrial process, illustrated in figure 8.

6. CONCLUSION

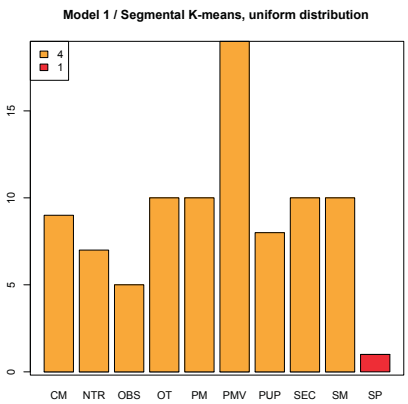
After testing randomness of our stochastic symbols generated by the synthetic model generator, we have successfully applied these two tests on three different HMM topologies. The first one, uses Aspin-Welch and the second one, uses Kolmogorov-Smirnov test. Given a set of observations sequences simulated by our synthetic model, we verified that the most relevant model had the *"goodness of fit"* i.e. how well model fits the set of observations sequences. In a statistical way, topology of model 2 is the best one. This corroborates results that model 2 is the one which comes closest to real industrial process in Vrignat et al. (2010). Thus, we specified our analysis from Roblès et al. (2011) paper. This criterion also shown that *Baum-Welch learning algorithm* with Forward Variable decoding gives best results with *normal distribution* of stochastic symbols. In our work on industrial breakdown prediction,



(a) HMM–Reference



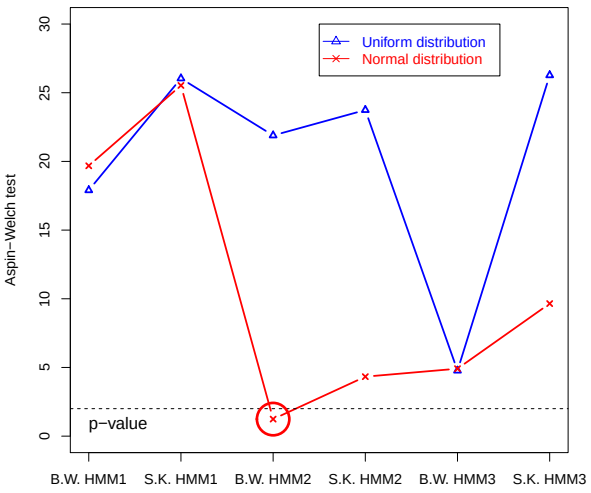
(b) Baum–Welch



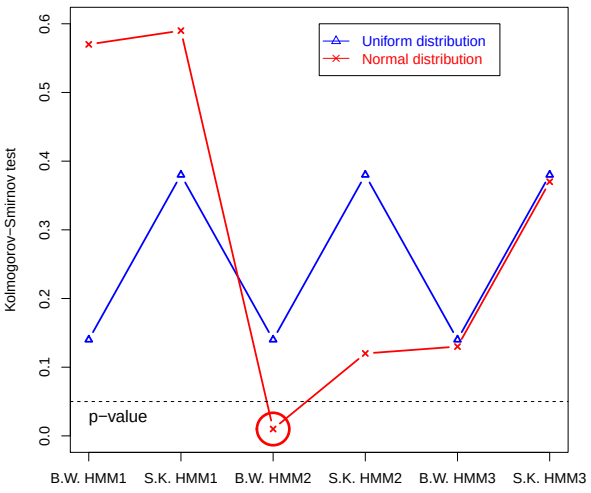
(c) Segmental K–means

Fig. 7. First sequence, using normal distribution

determining the best model is expected to reduce significantly failure rate in production and in the same time, to reduce power consumption. In further work, we will try



(a) Aspin-Welch test statistic



(b) Kolmogorov-Smirnov test statistic

Fig. 8. Statistical tests of Baum-Welch (B.W.) & Segmental K-means (S.K.) Algorithms

to test robustness of our synthetic model. Our researches, are able to validate a real choice of a model: topology, symbol,... without a priori knowledges on results.

REFERENCES

Baum, L.E., Petrie, T., Soules, G., and Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *The Annals of Mathematical Statistics*, 41(1), pp. 164–171.

Bercu, B. and Chafaï, D. (2007). *Modélisation stochastique et simulation - Cours et applications*. Collection Sciences Sup - Mathématiques appliquées pour le Master, Société de Mathématiques Appliquées et Industrielles (SMAI), éditions Dunod.

Drezner, Z., Turel, O., and Zerom, D. (2010). A modified kolmogorov-smirnov test for normality. *Communications in Statistics - Simulation and Computation*, 39, 693–704.

Fox, M., Ghallab, M., Infantes, G., and Long, D. (2006). Robot introspection through learned hidden markov models. *Artif. Intell.*, 170(2), 59–113.

Held, L., Rufibach, K., and Balabdaoui, F. (2010). A score regression approach to assess calibration of continuous probabilistic predictions. *Biometrics*, 66(4), 1295–1305.

Juang, B.H. and Rabiner, L.R. (1990). The segmental k-means algorithm for estimating parameters of hidden markov models. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 38(9).

Rabiner, L.R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceeding of the IEEE, 77(2) SIAM interdisciplinary journal*, 257–286.

Roblès, B., Avila, M., Duculty, F., Vrignat, P., and Kratz, F. (2011). Evaluation of relevance of stochastic parameters on hidden markov models. In *Advances in Safety, Reliability and Risk Management, European Safety and Reliability Conference (ESREL), ISBN : 978-0-415-68379-1*, 71. Taylor & Francis Group, Troyes, France.

Rukhin, A., Soto, J., Nechvatal, J., Barker, E., Leigh, S., Levenson, M., Banks, D., Heckert, A., Dray, J., Vo, S., Rukhin, A., Soto, J., Smid, M., Leigh, S., Vangel, M., Heckert, A., Dray, J., and Iii, L.E.B. (2010). A statistical test suite for random and pseudorandom number generators for cryptographic applications.

Steinebach, J. (2006). E. l. lehmann, j. p. romano: Testing statistical hypotheses. *Metrika*, 64, 255–256.

Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, 13(2), 260–269.

Vrignat, P. (2010). *Génération d'indicateurs de maintenance par une approche semiparamétrique et par une approche markovienne*. Ph.D. thesis, Université d'Orléans.

Vrignat, P., Avila, M., Duculty, F., and Kratz, F. (2010). Use of HMM for evaluation of maintenance activities. *IJAIS, International Journal of Adaptive and Innovative Systems, Vol. 1, Nos. 3/4*, 216–232.

Wang, J. and Schaalje, G.B. (2009). Model selection for linear mixed models using predictive criteria. *Communications in Statistics - Simulation and Computation 38*, 788–801.

Welch, B.L. (1947). The generalization of student's problem when several different population varlances are involved. *Biometrika*, 34(1-2), 28–35.

Welch, B.L. (1951). Welch's k-sample test. *Biometrika*, 38, 330–336.

Zille, V., Bérenguer, C., Grall, A., Despujols, A., and Lonchampt, J. (2007). Modelling and performance assessment of complex maintenance programs for multi-component systems. *ESREDA 32nd Seminar Proceedings, Alghero, Espagne*.