



Runtime Integration of Machine Learning and Simulation for Business Processes

Authors: Francesca Meneghello, Fondazione Bruno Kessler, Trento, Italy & Sapienza University, Rome, Italy

Chiara Di Francescomarino, University of Trento, Trento, Italy

Chiara Ghidini, Fondazione Bruno Kessler, Trento, Italy





Context: Business Process Simulation

Optimization

What-if scenarios

Analysis



Monitoring

Process improvement

Redesign process











Alternative approach: Generative models

- LSTM networks used to predict the activity of ● the next event in a case, its timestamp, and the resource/role associated to the event
- The approach has the ability to generate ۲ complete sequences from scratch by repeatedly predicting the subsequent event.



1 Discovering generative models from event logs: data-driven simulation vs deep learning, Manuel Camargo, Marlon Dumas, Oscar González-Rojas. (PeerJ Computer Science 2021)



Alternative approach: Generative models

- LSTM networks used to predict the activity of ulletthe next event in a case, its timestamp, and the resource/role associated to the event
- The approach has the ability to generate lacksquarecomplete sequences from scratch by repeatedly predicting the subsequent event.





Able to consider the features of the context/ scenario for prediction



Not suitable for what-if scenarios

Not global view, only on the single trace



1 Discovering generative models from event logs: data-driven simulation vs deep learning, Manuel Camargo, Marlon Dumas, Oscar González-Rojas. (PeerJ Computer Science 2021)



Hybrid Simulation: the Idea

Simulation model



Hybrid Simulation: the Idea

Simulation model







Hybrid Simulation: State of the Art



Learning Accurate Business Process Simulation Models from Event Logs via Automated Process Discovery and Deep Learning, Manuel Camargo, Marlon Dumas, Oscar González-Rojas. (CAISE 2022)

Simulated Log



Simulated Log

















Problems with Post-integration







Problems with Post-integration























Simulation model



Predictive model









Simulation model

Predictive model











Simulation model

Predictive model



RIMS: Runtime Integration of Machine Learning and Simulation for Business Processes





Predictive models for time perspective

Data Driven Process Simulation (DDPS) approach

RIMS receipt



- 1 Definition of the DDPS elements
- 2 Training of the predictive models
- 3 Integration and run simulation

Definition of the DDPS elements 1



Training predictive models 2



Time series to predict the trace start times

Predictive model for Waiting time:



- 1) Next activity
- 2) End timestamp of the current activity
- 3) Intercase features: WIP(work-in-progress) and resources' occupation



Predictive model for Processing Time:

- 1) Current activity
- 2) Start timestamp of the current activity
- 3) Intercase features: WIP(work-in-progress) and resources' occupation







Rojas. (CAISE 2022)













2



2























How does RIMS perform in terms of simulation quality compared to other techniques?



Evaluation: Procedure

- 25 simulations for each simulation approach
- MAE (Mean Absolute Error) of cycle times
- EMD (Earth Mover's Distance) of the normalized histograms of activity timestamps grouped by day and hour

Evaluation: Datasets

Log	Туре	#Traces	#Events	#Activity	Avg. Trace length
<u>Confidential1000</u>	Syn	800	21221	29	26.53
Confidential2000	Syn	1670	44373	29	26.57
<u>Cvs Pharmacy</u>	Syn	10000	103906	15	10.39
Purchasing Example	Syn	608	9119	21	15
<u>SynLoan</u>	Syn	2000	43164	25	39
Production	Real	225	4503	24	20
ConsultaDataMining	Real	954	4962	16	5.2
BPI12W	Real	8616	59302	6	6.88
BPI17W	Real	30276	240854	8	7.96

Evaluation RIMS: MAE metric

	DDPS				LSTM			Dsim			RIMS		
	Mean	CI_l	CI_u	Mean	CI_l	CI_u	Mean	CI_l	CI_u	Mean	CI_l		
Confidential 1000 Confidential 2000	55477 234520	51509 179570	59444 289471	32754 37344	32377 37063	33132 37625	36337	35796 7309	36878 8882	26950 7385	25683 6381		
Cvs Pharmacy	1061957	1061300	1062613	638665	637616	639714	137636	136104	139168	102950	101130	1	
PurchasingExample SynLoan	212/361 2295797	2109665 2295579	2145058 2296015	1446177	1446177	1463812	1675036	1635297	2287481 1714775	2104676 1278782	2082360 1271133	21 12	
Production ConsultaDataMining BPI12W BPI17W	312153 274786 667139 952502	307036 271536 665889 949997	317270 278035 668389 955007	128942 343520 316642 757742	117516 343422 309766 653898	140368 343618 323517 861585	83700 230921 597090 417641	73572 228349 594860 416631	93828 233494 599320 418650	105988 155994 555094 393749	91713 141180 550135 391442	1 1 5 3	

Evaluation RIMS: MAE metric

	DDPS				LSTM			Dsim		RIMS		
	Mean	CI_l	CI_u	Mean	CI_l	CI_u	Mean	CI_l	CI_u	Mean	CI_l	,
Confidential 1000	55477	51509	59444	32754	32377	33132	36337	35796	36878	26950	25683	;
Confidential 2000	234520	179570	289471	37344	37063	37625	8095	7309	8882	7385	6381	
Čvs Pharmacy	1061957	1061300	1062613	638665	637616	639714	137636	136104	139168	102950	101130) 1
PurchasingExample	2127361	2109665	2145058	2299619	2294319	2304919	2280717	2273954	2287481	2104676	2082360) 21
SynLoan	2295797	2295579	2296015	1446177	1446177	1463812	1675036	1635297	1714775	1278782	1271133	3 12
Production	312153	307036	317270	128942	117516	140368	83700	73572	93828	105988	91713	; 1
ConsultaDataMining	274786	271536	278035	343520	343422	343618	230921	228349	233494	155994	141180) 1
BPI12W	667139	665889	668389	316642	309766	323517	597090	594860	599320	555094	550135	; 5
BPI17W	952502	949997	955007	757742	653898	861585	417641	416631	418650	393749	391442	3

Evaluation RIMS: MAE metric

	DDPS				LSTM			Dsim			RIMS		
	Mean	CI_l	CI_u	Mean	CI_l	CI_u	Mean	CI_l	CI_u	Mean	CI_l		
Confidential 1000	55477	51509	59444	32754	32377	33132	36337	35796	36878	26950	25683		
Confidential 2000	234520	179570	289471	37344	37063	37625	8095	7309	8882	7385	6381		
Cvs Pharmacy	1061957	1061300	1062613	638665	637616	639714	137636	136104	139168	102950	101130		
PurchasingExample	2127361	2109665	2145058	2299619	2294319	2304919	2280717	2273954	2287481	2104676	2082360 2		
SynLoan	2295797	2295579	2296015	1446177	1446177	1463812	1675036	1635297	1714775	1278782	1271133 1		
Production	312153	307036	317270	128942	117516	140368	83700	73572	93828	105988	91713		
ConsultaDataMining	274786	271536	278035	343520	343422	343618	230921	228349	233494	155994	141180		
BPI12W	667139	665889	668389	316642	309766	323517	597090	594860	599320	555094	550135		
BPI12W*	694558	694083	695033	316642	309766	323517	310293	304608	315977	357940	351301		
BPI17W	952502	949997	955007	757742	653898	861585	417641	416631	418650	393749	391442		

Evaluation RIMS: EMD metric

	DDPS				LSTM			Dsim			RIMS		
	Mean	CI_l	CI_u	Mean	CI_l	CI_u	Mean	CI_l	CI_u	Mean	CI_l		
Confidential 1000	0.00115	0.00103	0.00127	0.01435	0.01434	0.01437	0.01269	0.01263	0.01274	0.01264	0.01257 0	0.0	
Confidential 2000	0.00109	0.00105	0.00113	0.00992	0.00991	0.00993	0.00897	0.00895	0.00899	0.00907	0.00905 0	0.0	
Cvs Pharmacy	0.00093	0.00092	0.00094	0.00237	0.00237	0.00237	0.00659	0.00658	0.00659	0.00672	0.00672 0	0.0	
PurchasingExample	0.00702	0.00699	0.00706	0.00525	0.00500	0.00550	0.00645	0.00634	0.00656	0.00856	0.00831 0	0.0	
SynLoan	0.00799	0.00798	0.00800	0.01926	0.01924	0.01928	0.00473	0.00469	0.00477	0.00321	0.00317 0	.0	
Production	0.01486	0.01438	0.01534	0.00715	0.00647	0.00784	0.00635	0.00589	0.00680	0.00523	0.00487 0	.0	
ConsultaDataMining	0.01573	0.01555	0.01591	0.03830	0.03820	0.03840	0.01233	0.01202	0.01265	0.00777	0.00739 0	.0	
BPI12W	0.00642	0.00640	0.00644	0.00766	0.00765	0.00768	0.00807	0.00806	0.00808	0.00805	0.00804 0	0.0	
BPI12W*	0.00468	0.00465	0.00470	0.00766	0.00765	0.00768	0.00799	0.00798	0.00799	0.00798	0.00798 0	0.0	

Evaluation RIMS: EMD metric

M		Dsim RIMS					
CI_l	CI_u	Mean	CI_l	CI_u	Mean	CI_l	
434 (991 (237 (500 (0.01437 0.00993 0.00237 0.00550	0.01269 0.00897 0.00659 0.00645	0.01263 0.00895 0.00658 0.00634	0.01274 0.00899 0.00659 0.00656	0.01264 0.00907 0.00672 0.00856	$0.01257 \ 0.00905 \ 0.00672 \ 0.00831 \ 0.00217 \ 0.00$).0).0).0
924 (547 (320 (765 (765 ().01928).00784).03840).00768).00768	0.00473 0.00635 0.01233 0.00807 0.00799	0.00469 0.00589 0.01202 0.00806 0.00798	0.00477 0.00680 0.01265 0.00808 0.00799	0.00321 0.00523 0.00777 0.00805 0.00798	0.00317 0 0.00487 0 0.00739 0 0.00804 0 0.00798 0).0).0).0).0

Evaluation RIMS: EMD metric

The End??

What about the queue??

What about the queue??

Enriched Simulated Log

Queue of Delivery Man Requests

What about the queue??

- Definition of the DDPS elements
- Training of the predictive models: 2
 - 2a Retrieve the queue feature from the log
 - ^{2b} Train the waiting time predictive model with the queue feature
- 3 Integration and run simulation

2a Retrieve the queue feature from the log

Retrieve the queue feature from the log 2a

2b

Train the waiting time predictive model with the queue feature

³ **RIMS⁺: Integration and run simulation**

Predict Waiting time

³ **RIMS⁺: Integration and run simulation**

Predict Waiting time

³ **RIMS⁺: Integration and run simulation**

How does RIMS+ improve the quality of simulation compared to approaches that do not use the queue feature?

Evaluation: Procedure

- 25 simulations for each simulation approach
- MAE (Mean Absolute Error) of cycle times
- EMD (Earth Mover's Distance) of the normalized histograms of activity timestamps grouped by day and hour

Evaluation: Datasets

Log	Туре	#Traces	#Events	#Activity	Avg. Trace length	Queue mea
<u>Confidential1000</u>	Syn	800	21221	29	26.53	9.91
Confidential2000	Syn	1670	44373	29	26.57	2.37
<u>Cvs Pharmacy</u>	Syn	10000	103906	15	10.39	59
Purchasing Example	Syn	608	9119	21	15	0.30
<u>SynLoan</u>	Syn	2000	43164	25	39	68.89
Production	Real	225	4503	24	20	1.17
ConsultaDataMining	Real	954	4962	16	5.2	0
BPI12W	Real	8616	59302	6	6.88	1115
BPI17W	Real	30276	240854	8	7.96	1646

Evaluation: Datasets

	Log	Туре	#Traces	#Events	#Activity	Avg. Trace length	Queue mea
	<u>Confidential1000</u>	Syn	800	21221	29	26.53	9.91
	<u>Confidential2000</u>	Syn	1670	44373	29	26.57	2.37
	<u>Cvs Pharmacy</u>	Syn	10000	103906	15	10.39	59
	Purchasing Example	Syn	608	9119	21	15	0.30
	<u>SynLoan</u>	Syn	2000	43164	25	39	68.89
r	Production	Real	225	4503	24	20	1.17
	ConsultaDataMining	Real	954	4962	16	5.2	0
	BPI12W	Real	8616	59302	6	6.88	1115
	BPI17W	Real	30276	240854	8	7.96	1646

Evaluation RIMS+ vs Best

			Best		
		Mean	CI_l	CI_u	Mean
	Confidential 1000	26950	25683	28217	15489
	Cvs Pharmacy	102950	101130	104771	45756
E	SynLoan	1278782	1271133	1286430	827635
MA	BPI12W	316642	309766	323517	454494
	BPI12W*	310293	304608	315977	301663
	BPI17W	393749	391442	396056	308038
	Confidential 1000	0.00115	0.00103	0.00127	0.01254
	Cvs Pharmacy	0.00093	0.00092	0.00094	0.00680
8	SynLoan	0.00321	0.00317	0.00325	0.00576
E	BPI12W	0.00642	0.00640	0.00644	0.00813
	BPI12W*	0.00468	0.00465	0.00470	0.00799
	BPI17W	0.00448	0.00446	0.00449	0.00847

RIMS+

- CI_u CI_l 16208 14771 53583 37930 821431 833839 452008 456980 292720 310606 299884 316193 0.01247 0.01260 0.00679 0.00681
- 0.00571 0.00582
- 0.00812 0.00814
- 0.00797 0.00801
- 0.00798 0.00895

- RIMS⁺outperforms the best approach in MAE
- RIMS⁺ worsens slightly in EMD

RIMS vs RIMS+: the impact of queue

6 Synthetic logs with different levels of queuing

Conclusions

- New hybrid simulation approach
- **RIMS** and **RIMS**⁺ outperform state of the art

Future work:

- Add the remaining perspectives
- Consider Resource Calendar
- Consider queue policies

THANK YOU! Questions, doubts, thoughts?

Contact me: <u>fmeneghello@fbk.eu</u> Process & Data Intelligence group website: <u>https://pdi.fbk.eu/</u>

THANK YOU! Questions, doubts, thoughts?

Contact me: <u>fmeneghello@fbk.eu</u> Process & Data Intelligence group website: <u>https://pdi.fbk.eu/</u>

