IOSR Journal of Engineering (IOSR JEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 10-17

Study of Technical Support for Mobile Process Mining

Mr. Dhiraj Rane[1], Dr. Mahendra Dhore[2]
[1]*MCA GHRIIT, Nagpur,*
[2]*Department of Electronics and Computer Science – RTMNU Nagpur*

Abstract: *A Process mining is relatively young research area that meets the gap between processes and various IT systems. Event logs are the primary sources for a processmining project and they are captured by different data sources including databases, ERPsystems, CRM systems, audit trails, mobile communication, hospital information systems, bank transaction logs,etc. The extracted knowledge from this log enable us to discover the actual process andexisting process model for further analysis, evaluation and continuous improvement intheir quality. This way, various process mining tools have been developed in the market.Nevertheless, there is a lack of sufficient and comprehensive evaluation frameworks thatassist users in selecting the right tool. This thesis provides the schematic view*

Keywords:*Process Mining, ProM, Simulation, Disco, MATLAB, NS2, RapidMiner*.

I. Introduction

Along with the rapid advancements in big data and cloud computing technologies, connection of everything is emphasized in many information systems. Thanks to the achievements of devices, infrastructure, and applications in mobile computing [1, 2], systems become more powerful and intelligent with the support of connection among devices, people, and business processes. Particularly, according to the recent research [3], mobile technology development has resulted in the creation of up to 1450,000 applications for smart phones in the last few years. More and more information systems rely on service-oriented processes in order to fit the continually changing business environment and to align business strategies with IT systems [4]. With strong interaction with people and social environments, these systems have a great impact in many areas such as health care [5, 6], exploiting indoor location [7], and other scenarios. As a result, it is becoming more and more valuable to deal with the connection among devices and interaction among people especially in the environment of the Internet of Everything (IoE).

Due to the flexible and scalable characteristics of service-oriented computing, more and more systems use web services composition to deal with the complexity of multisource data in mobile information systems. Business processes and associated services become the most significant supports for the connection of everything. They make functions and devices work as expected in well-organized systems. Achieving adaptiveness in process-based service composition is the key to improve efficiency and adaptiveness of mobile systems.

However, as both the scale and the variety of devices are expanding, the complexity of service implementation is increasing. To sum up, challenges exist in keeping the system process adaptive to the changing environment as the following points:

1) Process execution environment is changing: in the environment of IoE, as users, devices, and services are widely distributed, the execution of the process may be affected by changing device rules, connection situations, and event users' habits. As more complex rules are introduced with the devices, static processes always lack the consideration of execution environment, and they cannot handle the changing environment efficiently. For instance, in mobile systems, different versions of applications are used at the same time, which will make the processes in the server side suffer from errors if they cannot handle the changing orders of events.
2) The complexity of relationship in events and services is increasing: since types of devices are increasing, the relationship in events and services is getting more complicated. Current process-based service composition is not flexible enough to support the complex situations. As a result, approaches designed for application execution are usually incomplete and lacking necessary business consideration. For example, in a smart house application, when new devices like new models of air conditioners are introduced, new events and new connections will be introduced and the controlling process should be fixed accordingly in order to keep the devices and services work correctly.

In our previous work [8], the service composition based on process mining approach has been applied to a logistics cloud service platform which supports the users from different companies to customize their functional services. In the example case about the waybill transportation process, a suitable waybill-related

composite service is generalized to connect the information sensing devices like radio frequency identification (RFID), infrared sensors, global positioning system (GPS), and laser scanner. And it is proved that service composition based on process mining is suitable for the situation with indefinite requirements and without high performance demand of the result composite service. Considering the expanding scale and the variety of devices in mobile information systems, a process mining based service composition approach is proposed based on our previous work in this paper in order to improve the adaptiveness and efficiency of compositions.

Generally speaking, the main contributions in this paper can be summarized as follows:

(i) Firstly, to solve the problems above, process mining based service composition is proposed to produce adaptive service composition according to real execution information. A three-step framework is presented to cover the whole life cycle of service composition based on process mining.

(ii) Secondly, according to the framework, a set of models is put forward to support the holistic service composition approach which covers both the practical business and the execution effectiveness.

(iii) Then, to apply request-based logs in event-based process mining, a pre-processing algorithm is presented to transfer request-based logs to event-trace-based models so that the execution data can be used in process mining.

(iv) Last but not least, a scene-based service composition algorithm is presented in order to transfer the process mining results to service composition models which can be further used in service generation.

II. Related Work

Since process mining is a recent research subject, the focus tends in developing newtechniques and algorithms rather than developing a standard use case and attribute thatallows to compare various tools.According to Turner, Chris J., et al. [9] comparison of process mining tools was made inthe UK market, which elaborates the practice of business process mining along with themethod of analysis. The paper mentions types of business process mining, processmodels and process mining algorithms as a ground for comparing 7 process mining tools:Futura Reflect, Fluxicon, Comprehend, ARIS, BPM One, Iontas Focus Suite and FujitsuAutomated Process discovery Service.On the other hand, Irina Ailenei, Anne Rozinat, Albert Eckert and Wil M. P. Van der Aalst[10] mention: Unlike, assessing the quality of the discovered process model [11-12], a focuson the different functionalities related to the process mining use case context (role in thiscase) has been illustrated in the research work.A recent master's dissertation work provides an exploratory research on the use ofprocess mining tools by answering the questions [13]

- Why users choose a certain process mining tools?
- What are the criteria to choose those process mining tools?

The questions accordingly were addressed using in depth interviews and online surveys.In reference to the final findings: usability, visualization, integration and functionalitycriteria's including import functionality, export functionality and the presence of advancedfilter mechanisms by the tools was considered as an answer to research question.Nevertheless, unlike other related works, we adopt analytical methodology for capturingthe possible problems while mining a business processes in comparison to the operations(techniques) used to answer those problems. Furthermore, using process miningframework as a means to set the overall evaluation framework assures the dependabilityof the solution on process mining techniques. The thesis in general addresses issuesrelated to process mining tool selection that in turn enables to create transparent andanalysis based decision for the consumer.

III. Dedicated Tools

Among several available tools three of them are selected for further analysis at this stage.This includes: ProM 6, Fluxicon (Disco) and Celonis. For the commercial tools, theselection was based on availability of the tool for academic evaluation purpose.ProM (Process Mining) is toolset maintained by Eindhoven University of Technology,which collects prototypes developed in several research projects. The project is opensource and aims largely the academic and research group. The plugins added on demandenables to solve complex process exploration. ProM imports event logs compliant withthe MXML or XES formats and can load process model definitions in different standards.Some of the main features of ProM are: discovering the control-flow perspective of aprocess, social network analysis, analysing the resource and performance perspective of a process, discovering events based on decision rules and conformance checking with avariety of algorithms. ProM provides several export formats such as CSV and PNG.Disco is a commercial process mining tool developed by Fluxicon, which can run on topof Windows or Mac Desktops. It supports a wide range of event log import formatsincluding CSV, MS Excel, MXML, XES, FXL Disco Logs and DSC Disco project files.Some of the features include automated process discovery, animation of process maps,event log filtering with various parameters, project management and detailed statistics.Celonis is a commercial company that provides yet another software as a service processmining tool. Among the features for this tool include

Automated Integration of sourcedata, real-time surveillance of all business transactions, execution of process analyses,various filtering mechanisms and process reporting,

ProM

ProM is a generic open-source framework for implementing process mining tools in a standard environment. ProM (which is short for Process Mining framework) is an Open Source framework for process mining algorithms. ProM provides a platform to users and developers of the process mining algorithms that is easy to use and easy to extend. Our mission is to become the de facto standard process mining platform in the academic world by establishing an active, recognized community of contributors and users, and to create awareness for the power of process mining technology by promoting applications and industrial uptake. Our vision is to actively advance the state-of-the-art of process mining technology by developing methods that really work, by creating an open community, and by providing a stable and easily extensible platform, which optimally supports process mining.

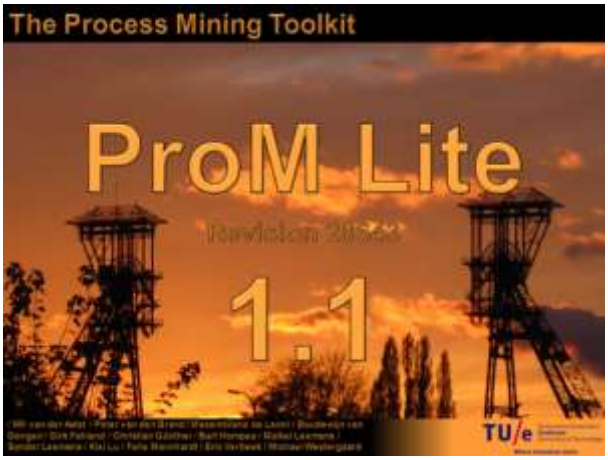

Figure 1:*Prom installation screen*

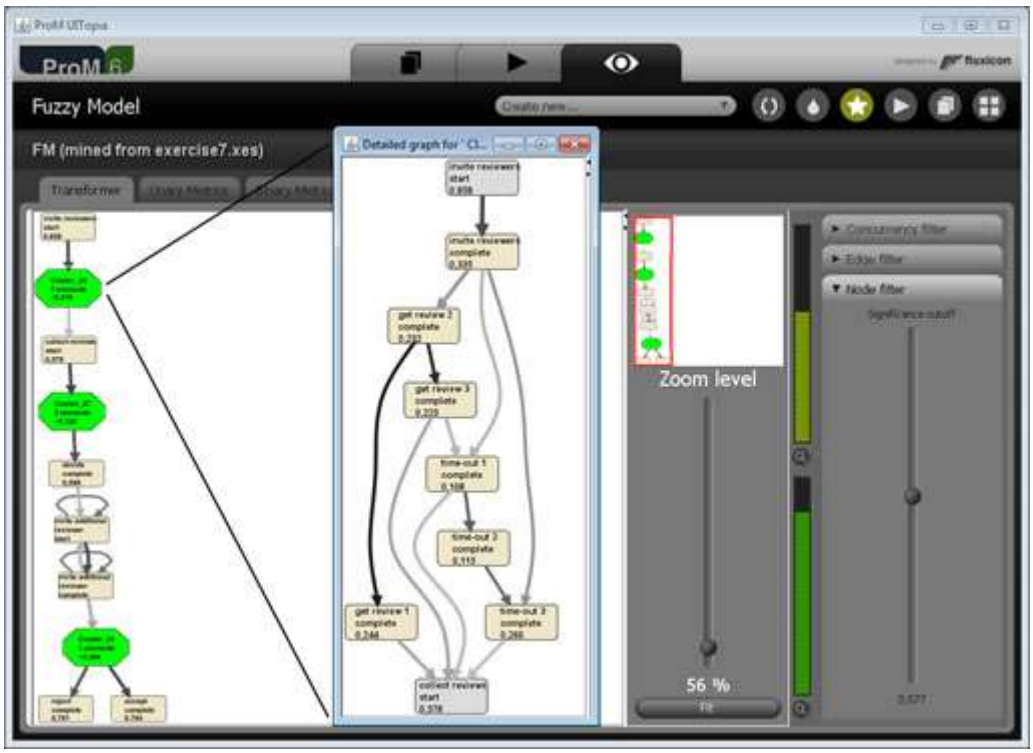

Figure 2:*Prom Process Analysis*

ProM for RapidMiner

ProM for RapidMiner enables the use of ProM plugins in RapidMiner. RapidMiner is a system which supports the design and documentation of an overall data mining process. As part of this, it offers a comprehensive set of operators for techniques of data mining, machine learning, and statistical methods. Furthermore, it contains many structures for defining the control-flow of the process.

Now the ProM framework and the RapidMiner data analysis solution are connected. As such any discovery, conformance, or extension algorithm of ProM can be used within a RapidMiner analysis process or a dedicated process mining analysis can be constructed. For example, a process can be constructed in which a log file is read, the process is discovered using the ILP miner, and finally the discovered Petri net is saved as a PNML file.

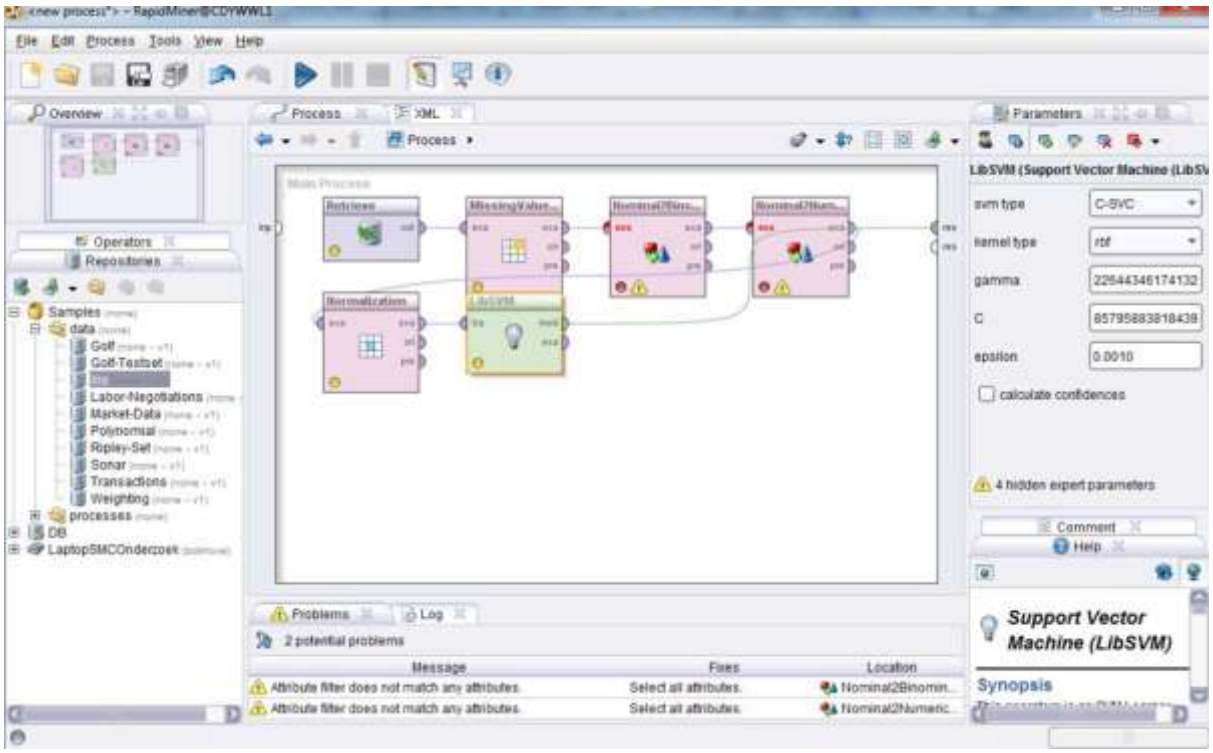

Figure 3:*RapidMiner Simulation*

XESame

XESame is an application that supports in the extraction of an event log from non-event log data sources. XESame is an application that supports in the extraction of an event log from non-event log data sources.OpenXES (pronounced as "open excess") is a reference implementation of the XES standard for storing and managing event log data. Application domains of the XES standard, and the OpenXES library, are manifold. They include, but are not limited to monitoring of process-aware information systems, process mining and general process analysis, and data mining.

XES is an open standard for storing and managing event log data. The OpenXES library is a reference implementation of that standard in Java™, which strives for strict XES compatibility, ease of development, and the best possible performance. OpenXES is released as open source / free software, under the terms of the GNU Lesser GPL (LGPL) license.

Both the XES standard and OpenXES have been designed with a tightly reduced, generally applicable core definition, with extensibility and generic applicability in mind. The XES standard provisions a generic extension system, which can be used to store data from arbitrary perspectives, or from specific application domains (e.g., for mapping corporate data models).

OpenXES

OpenXES (pronounced as "open excess") is a reference implementation of the XES standard for storing and managing event log data. Application domains of the XES standard, and the OpenXES library, are manifold. They include, but are not limited to monitoring of process-aware information systems, process mining and general process analysis, and data mining.

ProMimport
ProMimport is framework for the extraction of MXML-formatted logs from all kinds of popular information systems. ProMimport is framework for the extraction of MXML-formatted logs from all kinds of popular information systems. It is released under the terms of the revised BSD license, and requires a present installation of the Java Runtime Environment, version 1.5/5.0 or greater (Version 6.0 is recommended for Windows and Linux, 5.0 for Mac OS X).

ProM CPN Library
The ProM CPN library is a set of ML functions that supports the creation of MXML files from the simulation a Coloured Petri net. The ProM CPN library is a set of ML functions that supports the creation of MXML files from the simulation a Coloured Petri net. This library was created to facilitate experimenting with various process mining techniques/plug-ins

MXMLib
MXMLib is a Java library, used by ProMimport among others, when converting logs to the MXML format. MXMLib is a Java library, used by ProMimport among others, when converting logs to the MXML format. It is released under the terms of the revised BSD license, and can be incorporated in any program freely.
It requires a present installation of the Java Runtime Environment, version 1.5/5.0 or greater (Version 6.0 is recommended for Windows and Linux, 5.0 for Mac OS X).
Its features include:

- Efficient, straightforward composition and serialization of MXML documents.
- Efficient, straightforward composition and serialization of general XML documents.
- Simple, pervasive configuration framework for keeping the internal state of your application persistent.

Pentaho
Pentaho is a business intelligence (BI) software that provides data integration, OLAP services, reporting, information dashboards, data mining and extract, transform, load (ETL) capabilities. It is headquartered in Orlando, Florida. Pentaho was acquired by Hitachi Data Systems in 2015.On September 19, 2017, Pentaho became part of Hitachi Vantara, a new company that unifies the operations of Pentaho, Hitachi Data Systems and Hitachi Insight Group.
Pentaho offers an enterprise and community edition of the software. The enterprise software is obtained through an annual subscription and contains extra features and support not found in the community edition. Pentaho's core offering is frequently enhanced by add-on products, usually in the form of plug-ins, from the company and the broader community of users.

Disco
The revolutionary process mining technology in Disco can create beautiful and insightful process maps directly from your raw data, automatically. Pick your desired level of abstraction, choose from six process metric visualizations projected right on your map, and create filters directly from activities or paths. No matter how large or complex your process is — Disco will help you make sense of it.Sometimes you need to break out of static and detailed analysis views and see your process from a birds-eye perspective. With Disco you can create breathtaking animations, visualizing your process as it happened, right on your process map. Animation can help you to instantly spot bottlenecks where work is piling up.Is Raymond really the fastest in wrapping up support cases as he claims? The statistics view in Disco can answer all these questions, and more than you can imagine. Get an overview about your data from sleek, interactive charts, and drill down into detailed information about each activity, resource, and attribute value. It's all right there, at your fingertips

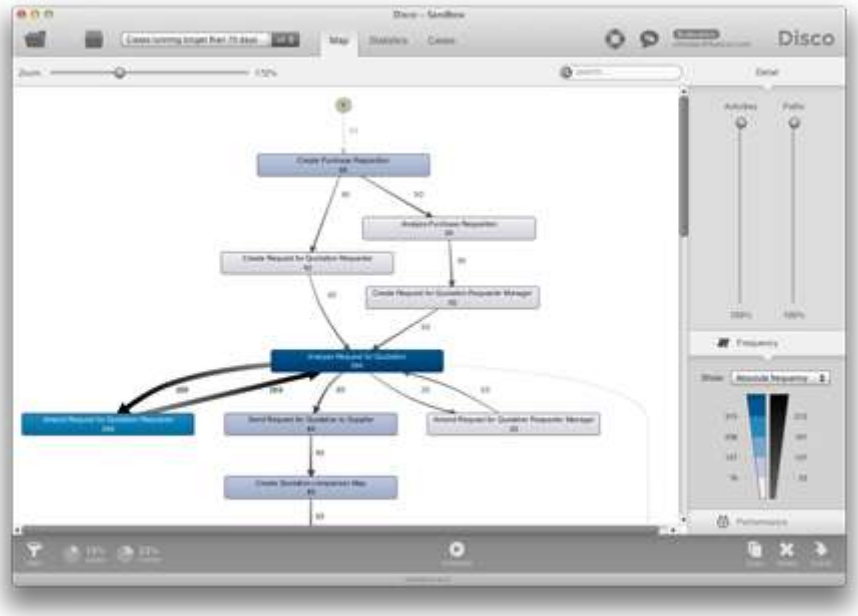

Figure 4:*Disco Process Analysis*

ProcessAnalyzer

QPR ProcessAnalyzer is the tool for turning event and transactional data into visual process intelligence. Read Introduction to QPR ProcessAnalyzer for an introduction of the software. Topics in this documentation are divided into user's roles: process analysts, developers and administrators.QPR ProcessAnalyzer is used via web browser. In QPR ProcessAnalyzer application you can perform various analyses on your process model. If you prefer standalone installation, you can use excel client for analyses.To create QPR ProcessAnalyzer models, you need to Extract data from source system(s), transform data to a format compatible with QPR ProcessAnalyzer and finally load data as process model(s). To perform these tasks, you can use T-SQL and QPR ProcessAnalyzer scripting commands. When you have created a model, end-users can analyze it easily by using QPR UI's ProcessAnalyzer application. If needed, there are plenty of tools to customize the end-user experience in QPR UI.

IV. Other Supporting Tools

NS2 Network Simulator

NS2 stands for Network Simulator Version 2. It is an open-source event-driven simulator designed specifically for research in computer communication networks.

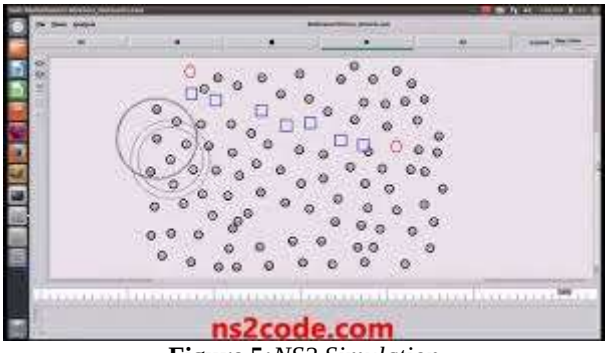

Figure 5:*NS2 Simulation*

Features of NS2

1. It is a discrete event simulator for networking research.
2. It provides substantial support to simulate bunch of protocols like TCP, FTP, UDP, https and DSR.
3. It simulates wired and wireless network.
4. It is primarily Unix based.

5. Uses TCL as its scripting language.
6. Otcl: Object oriented support
7. Tclcl: C++ and otcl linkage
8. Discrete event scheduler

Basic Architecture

NS2 consists of two key languages: C++ and Object-oriented Tool Command Language (OTcl). While the C++ defines the internal mechanism (i.e., a backend) of the simulation objects, the OTcl sets up simulation by assembling and configuring the objects as well as scheduling discrete events. The C++ and the OTcl are linked together using TclCL

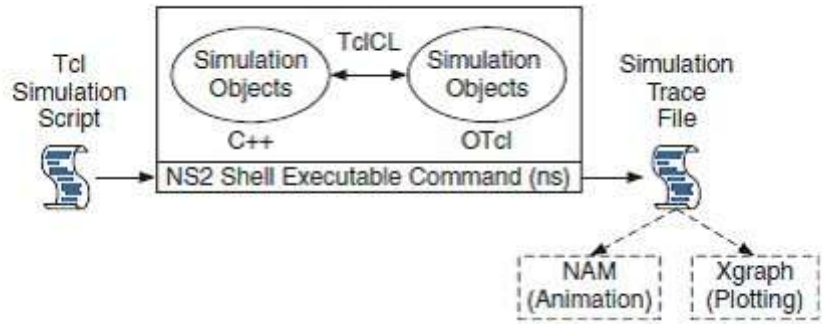

Figure 5:*NS2 Architecture*

MATLAB Simulation Toolbox

Wireless engineering teams use MATLAB® to reduce development time, eliminate design problems early, and streamline testing and verification.

- Prove algorithm and system design concepts with simulation and over-the-air signals
- Generate customizable waveforms to verify conformance to the latest 5G, LTE, and WLAN standards
- Create models using digital, RF, and antenna elements to explore and optimize system behavior
- Automatically generate HDL or C code for prototyping and implementation without hand-coding
- Create reusable golden reference models for iterative verification of wireless designs, prototypes, and implementations
- Automate analysis of large-scale field test data and visualize your simulation results

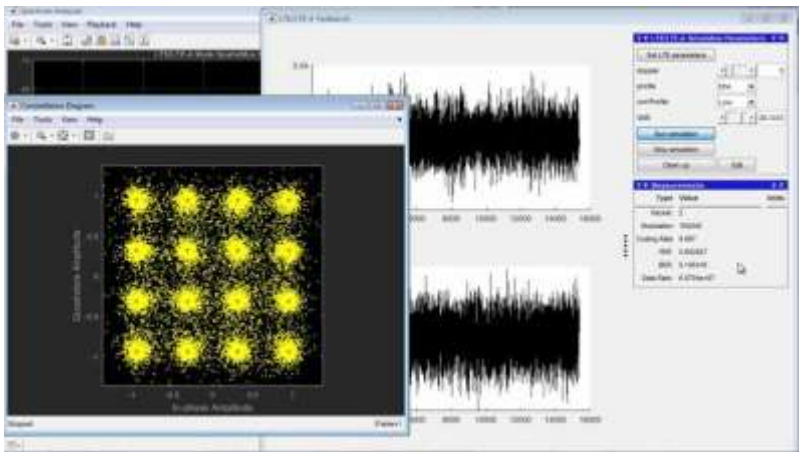

Figure 6:*MATLAB Wireless Simulation*

V. Conclusion

Although the research has reached its aim, there were some unavoidable limitations.First, we evaluated the tool using one specific scenario which is based on core techniquesof process mining tools as implemented by case studies. Therefore, it was not possibleto generalize the results. Second, the fact that we analyzed the tools using thedocumentation that is insufficient in availability so that the interpretation of the findingsmight be affected. Finally, the conclusions of the comparison might be different if theevaluation has been performed by expert users of the respective tools or by the toolvendors themselves.

References

[1]. H. Hoehle and V. Venkatesh, "Mobile application usability: conceptualization and instrument development," MIS Quarterly, vol. 39, no. 2, pp. 435–472, 2015. View at Google Scholar · View at Scopus

[2]. F. F. Ntawanga, A. P. Calitz, and L. Barnard, "A context-aware model to improve usability of information display on smartphone apps for emerging users," The African Journal of Information Systems, vol. 7, no. 4, p. 3, 2015. View at Google Scholar

[3]. V. V. S. M. Chintapalli, W. Tao, Z. Meng, K. Zhang, J. Kong, and Y. Ge, "A comparative study of spreadsheet applications on mobile devices," Mobile Information Systems, vol. 2016, Article ID 9816152, 10 pages, 2016. View at Publisher · View at Google Scholar · View at Scopus

[4]. H. Cai, C. Xie, L. Jiang, L. Fang, and C. Huang, "An ontology-based semantic configuration approach to constructing Data as a Service for enterprises," Enterprise Information Systems, vol. 10, no. 3, pp. 325–348, 2016. View at Publisher · View at Google Scholar · View at Scopus

[5]. B. Xu, L. D. Xu, H. Cai, C. Xie, J. Hu, and F. Bu, "Ubiquitous data accessing method in iot-based information system for emergency medical services," IEEE Transactions on Industrial Informatics, vol. 10, no. 2, pp. 1578–1586, 2014. View at Publisher · View at Google Scholar · View at Scopus

[6]. C.-W. Shen, C.-H. Hsu, C.-C. Chou, and T.-C. Tsai, "Toward a nationwide mobile-based public healthcare service system with wireless sensor networks," Mobile Information Systems, vol. 2016, Article ID 1287507, 11 pages, 2016. View at Publisher · View at Google Scholar · View at Scopus

[7]. H.-Y. Noh, J.-H. Lee, S.-W. Oh, K.-S. Hwang, and S.-B. Cho, "Exploiting indoor location and mobile information for context-awareness service," Information Processing & Management, vol. 48, no. 1, pp. 1–12, 2012. View at Publisher · View at Google Scholar · View at Scopus

[8]. Y. Li, H. Cai, C. Huang, and F. Bu, "Leveraging process mining on service events towards service composition," in Advances in Services Computing—9th Asia-Pacific Services Computing Conference (APSCC '15), pp. 195–209, 2015. View at Google Scholar

[9]. Turner, Chris J., et al. "Process mining: from theory to practice." Business ProcessManagement Journal 18.3 (2012): 493-512.

[10]. A.K. Alves de Medeiros and C.W. G¨unther. Process mining: Using CPN tools to create test logs for mining algorithms. Proceedings of the Sixth Workshop and Tutorial onPractical Use of Coloured Petri Nets and the CPN Tools, 2005.

[11]. S. Goedertier, D. Martens, J. Vanthienen, and B. Baesens. Robust Process Discovery with Artificial Negative Events. Journal of Machine Learning Research, 10:1305–1340, 2009.

[12]. [7] J. Munoz-Gama and J. Carmona. A Fresh Look at Precision in Process Conformance. In R. Hull, J. Mendling, and S. Tai, editors, Business Process Management (BPM 2010),volume 6336, pages 211–226, 2010.

[13]. Diederik Verstraete. Process mining in practice: comparative study of process miningsoftware (Masters Dissertation). Ghent University, Belgium, 2013