Provenance-based Root Cause Analysis for Revenue Leakage Detection: A Telecommunication Case Study

Wisam Abbasi[1][0000−0002−6901−1838], and Adel Taweel[1,2][0000−0003−0240−9857]

1 Computer Science, Birzeit University, Palestine. 2 Informatics, King's College London, UK
wisam.alabbasi@gmail.com
ataweel@birzeit.edu

Abstract. Revenue Assurance (RA) represents a top priority function for most of the telecommunication operators worldwide. Revenue leakage, if not prevented, depending on the severity of the leakage affecting their profitability and continuity, could cause a significant revenue loss of an operator. Detecting and preventing revenue leakage is a key process to assure telecom systems and processes efficiency, accuracy and effectiveness. There are two general revenue leakage detection approaches: big data analytics and rule-based. Both approaches seek to detect abnormal usage and profit trend behaviour and revenue leakage based on certain patterns or predefined rules, however both are mainly human-driven and fail to automatically debug and drill down for root causes of leakage anomalies and issues. In this work, a rule-based RA approach that deploys a provenance-based model is proposed. The model represents the workflow of critical RA functions enriched with contextual and semantic information that may detect critical leakage issues and generate potential leakage alerts. A query model is developed for the provenance model that can be applied over the captured data to automate, facilitate and improve the current process of root cause analysis of revenue leakages.

Keywords: Debugging · Provenance · Revenue Assurance · Root cause.

1 Introduction

The main responsibility of RA analysts is to manage and prevent revenue leakage based on RA methodology [8]. Current RA architectures support revenue leakage detection by applying a series of detective processes consisting of monitoring, summarization, auditing, and investigation [8]. But it is not an easy task to track back to the sources and root causes of a leakage issue manually due to the wide variety of rate plans, products, offers, campaigns, incidents, upgrades and millions or even billions of records in addition to the existence of tiered product plans and flat rates [3]. Therefore, automating the debugging and drill down process would greatly increase performance, ease the auditing process, save

operators revenues, provide better analytical experience, better management of data, more accurate reports and leads to an informed future decision making.

On the other hand, provenance is a global term refers to the creation history of an object and to the contextual information related to it [7]. Data Provenance has been applied to the field of computing by defining the origin of processes that have led to a specific state of data product within an information system such as databases and workflows [6]. Scientific research have been conducted in computer science for the purpose of data provenance application upon two major domains, science with the aim of information sharing and validation while preserving copyright and authority aspects [2] and business [1] to achieve data quality, reproducibility, auditability, validation, debugging, accountability, error backtracking, prediction, and forward tracking aspects [5].

Revenue leaks in telecom industry largely affect business profitability and continuity, and there is a critical need in the market for an RA analytics debugging tool as the debugging and drill down processes in current RA architectures are done manually. Therefore, we are proposing a new uniquely provenance data model for this domain, that would help revenue analysts audit their operators traffic in a better way, simplify error tracing back, better management of the data quality, and the provision of a historical record of data products.

2 The proposed approach

Provenance data represent semantic and contextual information related to the leakage issue or the telecom usage anomaly. These data are being created automatically based on the processes and sub-processes flow on the RA system. Each RA detective process consists of a number of processing steps and entities.

Figure 1 presents an overview of how the proposed provenance model works. Once an RA detective process starts execution the query model starts capturing semantic and contextual provenance information of each of its sub-processes from connected entities to each process, and store these information into data-oriented workflows as provenance diagrams in a graph database.

Entities that represent source nodes are associated with other entities to provide contextual information related to these nodes such as system logs, incidents and launched offers. The final result data item in the graphical workflow is associated with an entity named public holidays and events based on the date parameter to add more contextual information

Processes and entities in the provenance diagram are connected using relationships. The relationships are given the properties for backward tracing purpose. Semantic information represented by mapped attributes, and used filters if any are stored at the processing node level.

3 Running examples

The review of the existing RA systems suggests that they lack the provenance capturing capability to answer the questions of when, where, and why the issue

has been introduced and what reason(s) may have caused the revenue leakage issue. The proposed approach has been partially implemented and tested on two revenue leakage scenarios. Initial results show that it can automate the detection and root cause analysis of these scenarios.

Fake (i.e. false positives) revenue leakage alert: a great increase in voice calls duration with no additional increase in call counts or call charges has been used as an example of an abnormal behavior in the usage monitoring trends, since any additional voice usage must introduce additional fees or charges in the normal case, thus must be investigated by the RA team. The root cause for this behavior is that an offer was launched for local voice calls for prepaid subscribers to get 3 free minutes in each charged call.

Real (i.e. true positives) revenue leakage alert: a great missing of voice calls in one of the main RA system functions that connects the MSC source node representing the switch and the CCN source node representing the charging system has been used as an example of a real revenue loss as in figure 2 and

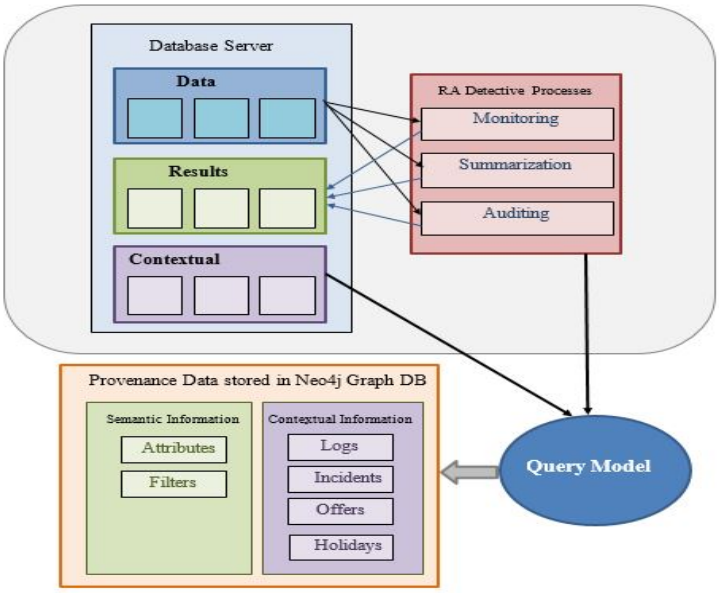

Fig. 1. Proposed provenance-based approach.

must be investigated immediately by RA team. The root cause for this issue is that calls charging has stopped due to a switch-charging disconnection occurred after an upgrade was done to MSC3 without adjusting its settings properly.

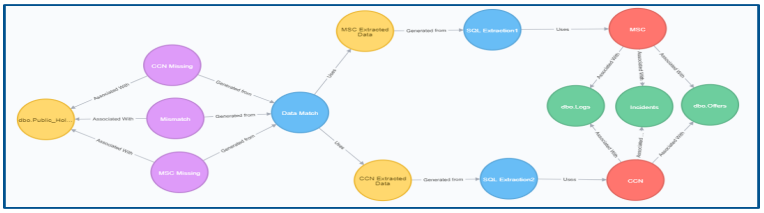

Fig. 2. Real revenue leakage alert scenario provenance graph.

4 Results

We have presented an approach that supports root cause analysis and drill down capabilities in current rules based RA systems. The approach starts by understanding the current state of RA function methodology, processes and approaches, then the current approach was improved through the deployment of logical data provenance and data workflows. For this stage, a preliminary evaluation was conducted to show the potential of the proposed approach and its plausibility on two scenarios, however for the next stage, the evaluation will include several scenarios and enlarged scope. The proposed model has been evaluated based on the accuracy and correctness of answering the questions of how, what, where, and why presented using the debugging and drill down concepts and it has proven its correctness and accuracy depending on the provenance data.

References

1. Curbera, Francisco, et al.: "Business provenancea technology to increase traceability of end-to-end operations.". In: " OTM Confederated International Conferences" On the Move to Meaningful Internet Systems". Springer, Heidelberg (2008).
2. Greenwood, Mark, et al.: "Provenance of e-science experiments-experience from bioinformatics.", 223–226 (2003)
3. revenue assurance how to stop bleeding and start leading, `https://clarity.sutherlandglobal.com/blog/accounting-minute/revenue-assurance-how-to-stop-bleeding-and-start-leading/`. Last accessed 7 Jan 2018
4. Global Telecom Revenue Assurance Survey 2013, `http://www.ey.com/Publication/vwLUAssets/Global\_telecoms\_revenue\_assurance\_survey\_2013/\$FILE/Global\_revenue\_assurance\_survey\_2013.pdf`. Last accessed 7 Jan 2018
5. Imran, Muhammad, and Helmut Hlavacs.: Provenance in the cloud: Why and how?. In: The Third International Conference on Cloud Computing, GRIDs, and Virtualization , pp. 106–112. CLOUD COMPUTING. (2012).
6. Moreau, Luc, et al.: "The provenance of electronic data." Communications of the ACM 51.4, 52–58 (2008)
7. Provenance, `https://en.wikipedia.org/wiki/Provenance`. Last accessed 7 Jan 2018
8. Revenue Assurance, `https://en.wikipedia.org/wiki/Revenue\_assurance`. Last accessed 7 Jan 2018