ETL Processes Security Modeling

Salma Dammak, University of Sfax, MIRACL-ISIMS, Sfax, Tunisia
Faiza Ghozzi, Sfax University, MIRACL-ISIMS, Sfax, Tunisia
Faiez Gargouri, University of Sfax, MIRACL-ISIMS, Sfax, Tunisia

ABSTRACT

The development of information systems (IS) in a secure environment or condition is a complex task that involves many additional basic security protocols, policies as well as industry standards on passwords, anti-virus programs, firewalls and data encryption. However, in traditional IS development lifecycles, security is either ignored or added as an afterthought, which does not assure the system complete security. So, it is necessary to give more importance to this issue and consider it as part of IS development process. In this context, the authors should guarantee the security of ETL (Extract, Transform, Load) processes, which are among the most critical and complex tasks during DW development project. In this study, security management is carried out for ETL processes by proposing a meta-model integrating the security concepts from the security requirements to the necessary preventive and / or corrective treatments. The proposed meta-model is validated with instantiation.

KEYWORDS

COSMIC, CVSS, ETL Processes, Measure, Meta-Model, Security, Vulnerability

1. INTRODUCTION

With the globalization of information exchanges, the winning companies become more and more dependent on the reactivity and safety of their information system (IS). The increase and complexity of the functionalities, to which the IS must respond immediately; make the realization of its conception and professional difficult and strategic. For this, reason, it is necessary to use models, methods, techniques and tools to ensure the easy and practical development of IS.

Due to the increasing complexity of data warehouses (DWs), a centralized and declarative management of metadata is essential for data warehouse administration, maintenance and usage. Data warehouses, more precisely the services they provide, play an increasingly important role in providing decision support, analytical reporting, ad-hoc queries and data mining for individuals, organizations and enterprises.

At the same time, DWs should function smoothly in a rather secure environment because of the increasing number of cyber-crimes. In these challenging conditions, the management of security risks is evolving and becoming more and more important for DWs that must be considered during their development. Indeed, security application is not easy at all levels of DWs development as it requires a lot of effort and continuous control.

Although it is difficult to have the right security information which covers both the security requirements and the current security state of an organization, a well-defined model has to be designed to describe the categories of the required information to develop secure DWs. The meta-
model definition may reduce the needed time and lower the security risk in the whole process of DWs security management.

As ETL processes are responsible for integrating, rearranging and consolidating a large volume of data through different sources as well as delivering data to the DW. Therefore, ETL should cover every phase, from designer’s security requirements to the design, implementation and testing steps. It must also ensure the security of data during extraction, processing and loading phases. However, works that proposed a secure modeling of ETL processes did not take into account ETL designer’s requirements. The integration of security requirements during ETL processes development allows satisfying ETL processes designer’s needs, improving the desired level of protection and enhancing decision making systems. For example, when we develop an ETL processes for any application domain (health systems or online bank, military system, etc.), we should protect sensitive data such as patient information, credit card number, critical value of turnover, politic information. Besides, a variety of vulnerabilities, affecting the ETL processes during their development or/and exploitation, can appear. If no preventive treatment is implemented, the detection of vulnerabilities leads to the loss of information, that may not be recoverable, and to waste of time and costs. Thus, an ETL designer, having no knowledge about security techniques, finds the integration of suitable solution to protect system hard and expensive in some cases.

To ensure the successful development of ETL processes, the below-mentioned problems should be solved:

- Translating designer’s security requirements to security solutions;
- Having an idea about vulnerabilities that can appear by anticipating them;
- Proposing both preventive and corrective treatment of security problems to be applied during their development and exploitation, respectively;
- Knowing which security problem to be treated firstly and its impact on the systems of ETL processes.

In this context, we suggested, in our previous work (Dammak et al., 2016), a GQVM (Goal Question Vulnerabilities Metric) approach which is an extension of the GQM approach (Basili, 1992). Indeed, GQVM translates designer’s security requirements to security goals of each step of developing ETL processes. In fact, to increase the security level, we anticipated a vulnerabilities catalog that can appear during ETL processes development. We also propose both the required preventive treatment to be integrated during ETL processes development and the necessary corrective treatment to be applied during exploitation phase. Moreover, we introduce a prioritization algorithm for vulnerabilities relying on two scores: the severity impact, measured according to the CVSS base score and the required treatment measured based on the COSMIC method. The GQVM approach helps ETL designer to develop secure ETL processes.

As previously mentioned, the use of designing models and methods ensures the easy development of DWs systems. Therefore, in this manuscript, we deeply investigate the security modeling of the defined GQVM approach to develop secure ETL processes. In fact, the introduced meta-model includes meta-class presenting the different concepts used in the GQVM approach: security goals, anticipated vulnerabilities, etc. To validate the proposed meta-model, we suggest object model to instantiate the different concepts introduced in the proposed meta-model.

The remainder of this paper is organized as follows: Section 2 defines security concepts and measurement standards. It also presents the related works concerning data warehouses (DW) and data integration security (DIS). Section 3 shows an overview about the GQVM approach and introduces our proposal of meta-model to develop secure ETL processes. Section 4 validates the meta-model and describes the developed framework of secure ETL processes development. In section 5, we discuss works dealing with security in data warehouse and data integration systems based on a comparative study. We also compare the limits and objectives of our research work to those of the existing studies. Section 6 is a brief conclusion.
Execution Management for Mobile Service-Oriented Environments
www.igi-global.com/article/execution-management-mobile-service-oriented/47037?camid=4v1a