Risk level assessment for data exchange applications in Digital Data Marketplaces

Lu Zhang, Arie Taal, Cees de Laat, and Paola Grosso

MultiScale Networked Systems (MNS) lab, University of Amsterdam, Science Park 904, Amsterdam, The Netherlands

Email: l.zhang2@uva.nl, A.Taal@uva.nl, delaat@uva.nl, p.grosso@uva.nl

The project Data Logistics for Logistics Data (DL4LD)[1] aims to facilitate secure and trust-worthy data sharing among Dutch logistic partners. The project pioneers the research on the concept of Digital Data Marketplaces (DDMs). A DDM is a membership organization that supports its consortium members to achieve a common goal through data and/or compute asset sharing.

Security and sovereignty are top concerns for such data federation applications. It is important for DDM customers to know how much security can be guaranteed by the DDM provider for its specific application. There is a basic need for techniques for addressing security assessment for data exchange procedures.

At ICT.Open, we will present an automatic, application-based, risk analysis-driven, collaborative risk assessment system for data exchange applications in DDMs. The general idea is that different data exchange applications may suffer from different vulnerabilities and have different threat models. In the implementation layer, various DDM providers also offer various sets of security countermeasures. All of the factors above contribute to different security levels guaranteed by a specific DDM provider for a concrete application. To achieve fairness and transparency, the evaluation system is collaborative. The risk assessment procedure is mainly performed by a trusted 3rd party, who is closely cooperating with DDM customers and providers. The trusted 3rd party estimates the risk level of all DDMs and provides the evaluation results to the DDM customers.

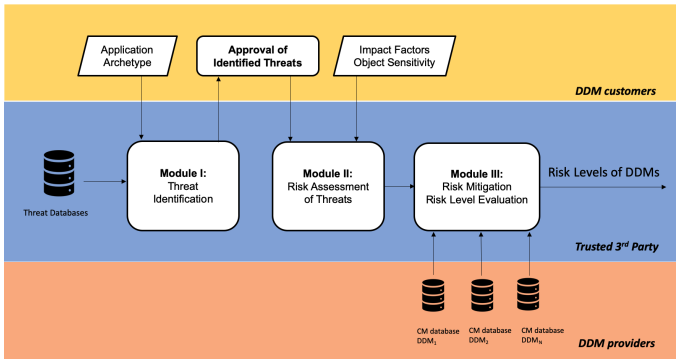

Figure 1: An application-based, risk analysis-driven collaborative risk assessment system of Digi

The architecture of the risk evaluation system is depicted in Figure 1. It is composed of 3 modules: i) *Threat Identification*; ii) *Risk Assessments of Threat*; iii)*Risk Mitigation and Risk Level Evaluation.*

DDM customers first feed their *application* to **Module I** operated by a trusted 3rd party. Then corresponding threats of the input *Application* are identified automatically. The generated threat list is sent to the DDM customer and each collaborating party checks the threat list. They sign the list if they agree, or go into a negotiation phase if some threats are missing. Only with all the signatures from the collaborating parties, the threat list can be further processed.

The signed threat list, as well as information like *impact factors* and *object sensitivity* , are fed into **Module II**. This module evaluates the application-based severity of each threat with the modified STRIDE/DREAD

[1]https://www.dl4ld.net/

Table 1: Risk attributes of modified DREAD Model

Risk Attributes	**Low (0)**	**Medium (5)**	**High (10)**
Damage Potential (DP)	Low Data Sensitivity	Medium Data Sensitivity	High Data Sensitivity
Accessibility (AC)	Only by consortium party member	By any involving party e.g. 3rd party	By outsiders
Skill Level (SL)	Advanced skills	Malware existing in Internet or using attack tools	Simple tools
Affected Users (AU)	One party member	Partial party members	All party members
Intrusion Detectability (ID)	Detectable without monitoring	Detectable by monitoring	Very hard to detect even by monitoring

model from Microsoft[1]. This model considers the possibility of an attack occurrence with 5 attributes and also the impact of each threat regarding the concrete *application*. Each attribute is scaled into 3 levels as High, Medium and Low and use parameters 0, 5, 10 to represent these three levels numerically. We modify those *risk attributes* to fit the DDM data exchange scenarios and the details are shown in Table 1.

Risk score RS represents the severity of an individual threat and is calculated as:

$$\begin{aligned} RS(t_i) = ImpactFactor(t_i) * \frac{1}{5}(DP(t_i) + AC(t_i) + SL(t_i) + AU(t_i) + ID(t_i)) \\ ImpactFactor(t_i) = ImpactFactor(C_n), for\, t_i \in C_n, \end{aligned} \tag{1}$$

where t_i denotes the ith threat and C_n denotes the nth threat category in STRIDE model.

Module III matches each threat with corresponding security countermeasures and computes the total risk level. The information about security countermeasures are provided in *DDM Countermeasure Databases* by DDM providers. The mappings can be one-to-one (1-1), one-to-multiple (1-N) or multiple to one (N-1). The originally 100% risk is divided into each threat based on its severity calculated in Module II. This indicates more dangerous threats contribute to higher risks. For each threat, the risk is reduced to different levels due to different countermeasures from DDMs. The *risk level* is finally quantified as the summation of the remaining risks of all the threats.

We will also present how the proposed system can be applied to the DL4LD use case. With the proposed system, the DDM customer can rank and select the most secure DDM for its own application. Additionally, we validate the resolution and parameter sensitivity for the algorithm, based on STRIDE/DREAD model, used in **Module II**. We first construct the DL4LD threat database by identifying the threats for all DL4LD data-exchange archetypes[2]. Then metrics, including *granularity*, *Kendall's Tau* and *Normalized Mean Square Error(NMSE)* are calculated based on simulation for both theoretical and DL4LD threat database. The simulation results will be presented at ICT.Open and they show that the algorithm is quite robust in terms of parameter sensitivity for those with same physical meaning. We are also able to provide guidance information for optimal parameter selection for those who would like to adopt our proposed method in other scenarios instead of DDMs.

References

[1] Miscosoft. Microsoft security development lifecycle (sdl). [Online]. Available: https://www.microsoft.com/en-us/securityengineering/sdl/

[2] L. Zhang, R. Cushing, L. Gommans, C. De Laat, and P. Grosso, "Modeling of collaboration archetypes in digital market places," *IEEE Access*, vol. 7, pp. 102 689–102 700, 2019.