Reachability-based impact as a measure for insiderness

Probst, Christian W.; Hansen, René Rydhof

Published in:
Journal of Wireless Mobile Networks, Ubiquitous Computing and Dependable Applications

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Reachability-based Impact as a Measure for Insiderness

Christian W. Probst1∗ and René Rydhof Hansen2
1 Technical University of Denmark, Denmark
cwpr@dtu.dk
2 Aalborg University, Denmark
rrh@cs.aau.dk

Abstract

Insider threats pose a difficult problem for many organisations. While organisations in principle would like to judge the risk posed by a specific insider threat, this is in general not possible. This limitation is caused partly by the lack of models for human behaviour, partly by restrictions on how much and what may be monitored, and by our inability to identify relevant features in large amounts of logged data. To overcome this, the notion of insiderness has been proposed, which measures the degree of access an actor has to a certain resource. We extend this notion with the concept of impact of an insider, and present different realisations of impact. The suggested approach results in readily usable techniques that allow to get a quick overview of potential insider threats based on locations and assets reachable by employees. We present several variations ranging from pure reachability to potential damage to assets causable by an insider.

Keywords: system models, insiderness, insider threats

1 Introduction

Countering insider threats is one of the big unsolved problems of organisational security. We lack the necessary insights to understand and explain human behaviour before the fact, and the tools we have at hand often fail even after the fact.

Mechanisms like logging and policy enforcement have been suggested to counter insider threats, and they probably are our best bet, but they come with a high price that make them difficult to use. In many parts of the world logging of employee actions is strictly regulated by law, and even if we were able to collect as much data as we wanted, it may be impossible to identify the important parts of that data. Policies come at the price of regulating ever bigger parts of employees’ work, and thus by adding more policies one risks to lower compliance with all or some policies [1].

What we lack are simple indicators for the risk that an organisation faces from insider threats. While this risk is based on individuals, evaluating their behaviour and the risk they pose, on a per person basis, is unrealistic of course. To overcome this problem, the notion of “insiderness” has been introduced [2], which is measured as a function of, e.g., access to and knowledge of potential targets and security measures. We lift this notion to the personal level based on impact, as measured in terms of what assets can be accessed and the possible damage to these assets.

In this work, we define several different notions of insiderness, allowing for a risk assessment that is better suited to the specific needs of a given organisation. The goal of these metrics for insiderness is to provide organisations with basic and very quick overview of the organisations security stance with
reachability-based impact as a measure for insiderness. Even better, the results can be used to evaluate and predict the risk of collaborating insiders in an easy manner.

In order to be able to scale our approach to even the largest of organisations, we have been inspired in our approach by work aiming at reducing complex properties by simple metrics, the primary example being the maintainability and performance of large software projects [3]. In a next phase we will evaluate the metrics introduced here on real world settings.

The rest of this article is structured as follows. We begin with a survey of related work in Section 2, including a brief discussion of the notion of insiderness as defined by Bishop et al. [2]. After this we introduce, in Section 3, a series of new concepts of insiderness that are all orthogonal to the work in [2]. In particular, our concepts facilitate fast techniques for computing rough measures for potential vulnerability to insider threats. Finally, Section 4 discusses our results and outlines future work.

2 Related Work

The notion of “insider” has been thoroughly re-examined and discussed by Bishop et al. [2], and arguing that instead of considering insider threats in a binary manner, i.e., inside(r) vs. outside(r), we should think of the degree of insiderness of a threat. The degree of insiderness is then considered as a function of access, knowledge, and, to some extent, trust, although it can be argued that trust is subsumed by the other two factors.

Bishop et al. further refine their definition of insider attacks as attacks that occur by exploiting gaps between ideal, or “oracle”, security policies and feasible (implementable) security policies. These ideas are then examined both in the context of the Unifying Policy Hierarchy [4] and the Attribute-Based Group Access Control Model (ABGAC) [5].

The measurement and prediction aspect of our work is inspired by metrics used for judging the quality, maintainability, and performance of big software development projects [3, 6]. Here the goal is to develop simple indicators for quality, which are independent from low level, detailed structures in the project and software. Since the metrics developed in this work are lightweight, they can be applied continuously.

In [7] a graph based model for assessing insider threats is presented, that is very similar in structure to the model we present in this paper. It also shares many of the same goals and generic approaches. However, the models presented in [7] do not have an underlying semantics, whereas the models presented in this paper are formally rooted in the Klaim family of process calculi [8, 9]. Furthermore, the notions discussed in this paper are mostly straightforward applications of reachability allowing for efficient implementation using standard algorithms.

3 Notions of Insiderness

In this section we introduce our different notions of insiderness. As already mentioned, the underlying idea is inspired by the work of Bishop et al. [2]. However, our approach is orthogonal to that work and of a more practical nature, with the explicit goal of deriving implementable metrics that scale to very large organisations.

Based on previous work on system models and formal static analyses of these models [2, 11] we treat locations and assets as resources with respect to insider threats. This leads to immediate notions of insiderness: “reachable locations” and “accessible assets”. Both these notions are easily formalised and represented in the system models mentioned above and can automatically be computed with respect to the access control specification of an organisation. Furthermore, it is straightforward to extend both
the system model and the notions of insiderness with quantitative concepts such as time of “difficulty to access”.

The metrics for insiderness we present in this section are rooted in an organisational model based on the ExASyM modelling formalism [11]. ExASyM models are graph-based representations of organisational structures annotated with access control specifications. We have in earlier work developed a number of algorithms that identify, for given actors, which nodes (in the graph underlying the model) they can reach in the organisation model based on their identity and known keys and given an access control specification.

In the following, we use this formalisation to define a number of metrics in terms of the set of reachable locations for a given actor, yielding a simple but flexible foundation for easily defining and investigating a wide spectrum of metrics. Based on the nodes in this set, we compute different indicators for the degree of insiderness of an actor. For all these metrics it holds that higher values represent a higher risk, or at least more substantial access to resources.

In addition to the above, actor-based metrics, we also outline how metrics based on more sophisticated approaches can be developed and how simple risk assessment can be performed based on the metrics defined below. A specific risk assessment could be, for example, to judge the impact of a set of insiders collaborating. Due to the simplicity of our approach, stemming from its foundation on sets of locations, risk assessment can be performed using straightforward union of sets and the same metrics used for individual actors. The obtained measures can also be used to group insiders into equivalence classes with respect to their level of insiderness.

The biggest challenge with the original notion of insiderness is, as discussed above, the difficulty of computing it. On the other hand, once it has been obtained, it is a valuable tool on a fine-grained level. The goal in designing our measures of insiderness was to derive metrics that are extremely easy to obtain and apply. While the results need to be interpreted carefully, we believe that the computed values are useful as initial indicators guiding more detailed risk assessment [12].

### 3.1 System Model

Recently, several system models have been introduced to model organisations and analyse them for vulnerabilities including, but not limited to, those posed by insider threats. Examples of such models include Portunes [13], ExASyM [11], and ANKH [14]. The former two have formal semantics based on (variants) of the Klaim process calculus [8, 9], while the latter uses the actor-network theory of Bruno Latour [15, 16] as a foundation.

We base our work on ExASyM and treat an organisational model as a graph:

$$\mathcal{M} := (\mathcal{N}, \mathcal{E}, \mathcal{D}, \mathcal{A}, \mathcal{AC})$$

where \(\mathcal{N}\) is a set of nodes representing locations in the organisation being modelled, \(\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}\) is a set of directed edges representing connections between locations, \(\mathcal{D}\) is a set of data items, \(\mathcal{A}\) is the set of actors in the system, and \(\mathcal{AC} \subseteq \mathcal{N} \times 2^{(\mathcal{A} \cup \mathcal{D} \cup \mathcal{N}) \times 2^{\text{actions}}}

\(\text{actions}\) represents access control specifications\(^1\) where the set of possible actions, denoted actions, depends on the specific system that is modelled. We therefore do not specify it further here.

Figures 1 and 2 show the blueprint of a small organisation and the associated ExASyM graph. As can be seen, there exists a direct mapping of locations in the blueprint to nodes in the graph, and from connections to edges. Note that the nodes in the system model in Figure 2 are annotated with access control specifications, e.g., the node representing the hallway, denoted HALL, has an access control

\[^1\text{We assume } \mathcal{N}, \mathcal{E}, \mathcal{D}, \text{and } \mathcal{A} \text{ to contain unique elements. In this article we ignore actions, access control, and other additions to the model. A discussion of actions in the ExASyM model can be found in [11].}\]
specification, ‘∗: m’, stating that all actors (indicated by ‘∗’) are allowed to move into that node (indicated by ‘m’). We do not here go into further details with the graphical representation, instead we refer to [11] for a detailed explanation and semantics of the system model.

The system model is used in the algorithm reachable, which takes as argument an actor with a certain knowledge representing, e.g., keys available to the actor and a start location, and returns the subset of $N$ that the actor can reach using this knowledge:

$$reachable: \mathcal{A} \times 2^\mathcal{D} \times N \rightarrow 2^N$$ (1)

$$R(a, D_a, loc) := reachable(a, D_a, loc)$$ (2)

where $a \in \mathcal{A}$ is an actor, $D_a \subseteq \mathcal{D}$ is a set of data items (known by the actors), and $loc \in N$ is a location. Thus $R(a, D_a, loc)$ is the set of locations (nodes) that are reachable by actor $a$ from location $loc$ by using only data, e.g., keys and access codes, in $D_a$. We omit (part of the) indices to $R$ wherever they are clear from context.

### 3.2 Actor-based Insiderness

The first set of notions of insiderness we present is based on actors, and how they have access to parts of the organisations and its assets. Starting from the simplistic notion of reachability, we define several notions that try to estimate the possible impact of an insider attack launched by an actor.

#### 3.2.1 Reachability

The simplest concept of insiderness is that of reachable locations. It computes for a given entity how big a part of an organisation this entity has access to. The reachability-based insiderness is measured as
percentage of locations reachable. In this simplistic version, reachability-based insiderness is computed per actor $a$, and assumes that the actor starts at a given location $loc$:

$$I_R(a) := \frac{\#R(a, \mathcal{D}_a, loc)}{\#\mathcal{N}}$$

As we will argue later, this is not necessarily the most useful level of abstraction, but it may rather be useful to compute properties of groups of insiders based on a set of credentials shared between them. The computation of values for individual actors then becomes a special case, where the credentials happen to be the ones owned by that actor.

### 3.2.2 Weighted Reachability

Obviously, not all locations in an organisation are equally important. An insiderness level of, e.g., 50% may be considered critical if the reachable locations are sensitive, while it may be considered insignificant, if it only comprises public or shared locations.

We now extend the notion of reachability-based insiderness by a weight that for each locations indicates an importance. This weight is given by a weight function $w: \mathcal{N} \rightarrow \mathbb{R}_{\geq 0}$ that maps a node to some numerical, non-negative representation of its weight and thereby its sensitivity. The idea is that larger numbers represent higher importance.

The insiderness based on weighted reachability is then computed by adding the weights for all nodes.
reachable by an actor $a$ as computed by

$$I_W(a) := \frac{\sum_{r \in R(a, D_a, loc)} w(r)}{\sum_{n \in N} w(n)}$$

(4)

Using the constant weight function $w(n) = 1$ for all nodes, this metric for insiderness is the same as the one based on reachability.

### 3.2.3 Accessible Assets

Both metrics we have introduced so far are merely based on the locations an insider can reach; here the reachability based on identity and knowledge of the actor is the important property. Another measure for the importance of a location is the possible impact of reaching that location. In the spirit of the other measures suggested in this paper we aim for an easily computed metric. To realise this we abstract the impact of reaching a location by considering the value of assets accessible at that location.

We now extend the notion of reachability-based insiderness by considering these accessible assets at each location. This basically means that we, as before, identify all locations reachable by an actor, add the values of assets at the reachable locations, and then relate this sum to the sum of values of assets accessible for all locations:

$$I_A(a) := \frac{\sum_{r \in R(a, D_a, loc)} \sum_{av \in assets(r, a, D_a)} value(av)}{\sum_{n \in N} \sum_{v \in assets(n)} value(v)}$$

(5)

The function $assets$ maps a location to a set of all assets accessible by the given actor with a set of keys at that location, and $value$ maps an asset to a numerical, non-negative value representing the asset’s value.

A simplified version of the accessible-assets insiderness would simply count the number of assets accessible by an actor. As in the case of weighted-reachability insiderness vs. reachability insiderness, this would be realized by setting the value of assets to 1 for all assets.

### 3.3 Location-based Insiderness

We now consider the insiderness of locations in the organisation. To compute this location-based insiderness we first identify all paths through a location, and then compute the sum of all assets that are accessible at the goal location of these path. The result basically represents the value of all assets that are accessible from nodes that can be reached from this location on paths that must lead through this location. In the example shown in Figure 1, the reception, the face recognition, and the hallway all would have a high location-based insiderness, since all attack paths will lead through them.

On the system model graph the value for this kind of insiderness insiderness is easily compute by considering for each node the subgraph that a location dominates. In computing this value we also must consider the other domains connected to a location in the subgraph. In the example’s system model graph it is clear that the nodes for the reception, the face recognition, and the hallway dominate the rest of the model for the physical connections. At the reception, for example, we then also need to consider the assets that can be accessed using the computer network.
3.4 Psycho-social Indicators

A more controversial, but potentially very powerful, measure of insiderness is the use of certain *psycho-social indicators* that may (or may not) indicate a person's likelihood, or willingness, to perform an (insider) attack as discussed by Bishop *et al.* [2]. It is important to stress that, currently, the statistical validity of these indicators has not been confirmed in sufficiently robust and credible experiments.

Nevertheless, even if these indicators turn out to be too imprecise or error-prone for predicting actions of individuals, they can still be useful for organisations both to raise awareness about these issues within the organisation as well as to determine if the *psycho-social security posture*, for lack of a better term, of the organisation has a psycho-social bias, *i.e.*, makes unintended assumptions about the psychological profile of typical insider attackers or non-attacking employees and thereby is left open to attack.

We leave it for future work to determine exactly which indicators to take into account and how to weight the different indicators.

3.5 Adding other Factors

In the previous sections we have seen a number of factors that can be used to refine the simple notion defining insiderness based on reachability. In principle we can continue adding more and more nuances that relate to insiders and the potential threat they pose for an organisation.

Some further interesting factors include the *time* needed to reach a certain location, the *risk of detection* while performing a certain action, or the *difficulty* of accessing a resource of a given type.

As future work, and as suggested by an anonymous referee, we conjecture it would also be possible to add factors that take into account how likely a given person is to intentionally exceed his/her granted access level, *e.g.*, by exploiting security vulnerabilities or through social engineering. This would enable more precise modelling of insiders that are willing to plan and perform active attacks. These factors could also take an employee’s job function into account to provide even finer grained modelling and analysis.

By now it should it be obvious, how such measures can be added to our approach. We are currently in the process of adding some of them to ExASyM to make them available for analyses.

3.6 Risk Assessment

One of the biggest challenges related to insider threats is to judge what influence an insider attack will have on the organisation [12]. Often, in-depth analyses will be considered too unreliable and too cumbersome. Interestingly this is the case for small organisations, due to lack of resources, as well as for large organisations, due to complex structures.

Using the techniques presented in this section, a rough measure for the risk associated with a certain actor or a group of actors having certain knowledge can be easily computed, as can the potential contribution of a certain location to the impact of an attack. Even more importantly, also the effect of collaborating insiders, or of an insider obtaining credentials of another insider, can be judged. Similarly, the techniques may be used to analyse the potential impact of specific people or job functions collaborating, *e.g.*, a technician with physical access to the disks of an otherwise restricted server and an accountant with specific knowledge of high-value content on the disks. Such modelling and analysis could lead to increased monitoring/logging of certain events and locations or policies that directly target specific (combinations of) job functions.

To perform this risk assessment one computes the set of reachable locations for each actor. By applying the metrics defined above to the union of the reachable locations of two or more actors, one obtains a measure for the insiderness of the group of actors. This approach is simple and straightforward and can easily be adapted by organisations of all sizes.
Figure 3: The online automaton for predicting locations of actors based on observed events in the system. The events on the edges are those actions in the system that are logged. The locations in the nodes of the automaton are log-equivalent locations; these locations can be reached from one of the other locations without causing a new log entry.

This approach to risk assessment supports also the computation of “dependent” insiderness, similar to dependent probabilities. Independent of the used metric one can interpret the insiderness resulting from one insider A obtaining credentials from another insider B as being dependent on B’s credentials. Using this information one can build trees of dependent insiderness impacts. This information can be valuable in understanding and explaining insider attacks.

3.7 Online Surveillance

In earlier work we presented an approach to transform the system model and its access control specification into a surveillance automaton, which is able to predict locations where an actor might be located [17]. This prediction is based on events observed in the system by means of, e.g., logging mechanisms, and on the notion of log-equivalent actions. Log-equivalent actions are those actions that we know are unobservable in a system, e.g., opening a door with a regular key. The automaton for the system model from Figure 2 is shown in Figure 3.

This kind of abstraction is ideally suited to combine the different kinds of insiderness presented. It allows to give an approximation of which actors are in which areas of the system. Using the insiderness of locations and the insiderness of the actors, it allows to perform a quick assessment of whether a location is under the risk of attack or not. This approach is especially suitable for locations with a potentially high impact in case of an insider attack.
4 Conclusion

Insiderness is a valuable tool in judging an organisation’s latent vulnerability to insider threats. In this work we pick up the original definition and extend it with an orthogonal notion of impact. This impact is based on reachability of locations and thereby assets that may be damaged.

Based on the original work of Bishop et al. we have introduced a number of metrics to measure the insiderness of actors with respect to an organisation. The metrics presented here are all based on concept of reachability, e.g., the set of nodes reachable by an actor. Using the set of reachable nodes we compute three metrics: reachability insiderness and weighted-reachability insiderness are purely location based, while accessible-assets insiderness also takes the potential damage to assets into account.

At the same time it is clear that the presented metrics are a very crude approximation of a value that in reality is not computable: the threat posed by an insider. This threat depends on too many factors that are neither predictable nor measurable. As we have stressed several times, the metrics have been designed with ease of application in mind, and we believe they are indeed straightforward to apply. The introduced metrics are easy to compute, they are modular, and are easy to interpret. Combined, these properties make them applicable at different sizes of organisations, and useful for a continuous surveillance of potential insider threats. Due to their simplicity they should also be easy to visualize, an important aspect in communicating and rationalizing the analysis results.

We are currently working in several directions. On the computational level we investigate how to add the factors mentioned above, especially time, risk of detection, and difficulty of accessing resources. These properties partly break the ease of performing risk assessment, since they are not modular in the same way as reachability, and they are not properties of the locations. On the organisational level we investigate the addition of psycho-social indicators, and here we face similar problems as just described (besides the problem of obtaining reliable input on what to add). Last but not least we work on case studies to show the applicability of the measures in real-life scenarios. Especially applications to more technical areas such as e.g., cloud computing seem promising [18].

5 Acknowledgments

The authors would like to thank the anonymous referees for numerous helpful and constructive suggestions. Part of the research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 318003 (TREsPASS).

References


Author Biography

**Christian W. Probst** is an Associate Professor in the Department of Applied Mathematics and Computer Science at the Technical University of Denmark, where he works in the section for Language-Based Technologies. The motivation behind Christian’s research is to realize systems with guaranteed properties. An important aspect of his work are questions related to safety and security properties, most notably insider threats. He is the creator of ExASyM, the extendable, analysable system model, which supports the identification of insider threats in organisations. Christian has co-organized cross-disciplinary workshops on insider threats and has co-edited a book on the topic.

**René Rydhof Hansen** is an Associate Professor at the Department of Computer Science, Aalborg University, Denmark. His research interests include security, static analysis, software verification and validation, real-time systems, and programming languages.