HAL
open science

Using the Conflicting Incentives Risk Analysis Method

Lisa Rajbhandari, Einar Snekkenes

To cite this version:

Lisa Rajbhandari, Einar Snekkenes. Using the Conflicting Incentives Risk Analysis Method. 28th Security and Privacy Protection in Information Processing Systems (SEC), Jul 2013, Auckland, New Zealand. pp.315-329, 10.1007/978-3-642-39218-4_24 . hal-01463835

HAL Id: hal-01463835

https://inria.hal.science/hal-01463835

Submitted on 9 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Using the Conflicting Incentives Risk Analysis Method

Lisa Rajbhandari and Einar Snekkenes
{firstname.lastname}@hig.no

Norwegian Information Security Laboratory, Gjøvik University College, Norway

Abstract. Risk is usually expressed as a combination of likelihood and consequence but obtaining credible likelihood estimates is difficult. The Conflicting Incentives Risk Analysis (CIRA) method uses an alternative notion of risk. In CIRA, risk is modeled in terms of conflicting incentives between the risk owner and other stakeholders in regards to the execution of actions. However, very little has been published regarding how CIRA performs in non-trivial settings. This paper addresses this issue by applying CIRA to an Identity Management System (IdMS) similar to the eGovernment IdMS of Norway. To reduce sensitivity and confidentiality issues the study uses the Case Study Role Play (CSRP) method. In CSRP, data is collected from the individuals playing the role of fictitious characters rather than from an operational setting. The study highlights several risk issues and has helped in identifying areas where CIRA can be improved.

Keywords: Risk analysis, risk, privacy, conflicting incentives

1 Introduction

Risk is usually expressed as a combination of likelihood and consequence but obtaining credible likelihood estimates is difficult. Thus, there is a need to improve the predictability and the coverage of the risk identification process. This challenge is a consequence of limited availability of representative historic data relevant for new and emerging systems. Besides, people are not well calibrated at estimating probabilities [20]. Furthermore, to improve the efficiency of the identification process, there is a need to identify issues that are key to risk discovery, and avoid activities that shed little or no light on potential problem areas. The Conflicting Incentives Risk Analysis (CIRA) [19] method addresses these issues by using an alternative notion of risk. In CIRA, risk is modeled in terms of conflicting incentives between the risk owner and other stakeholders in regards to the execution of actions. However, little evidence exists to suggest that CIRA is feasible to analyze risk in non-trivial settings.

In this paper, we explore to what extent CIRA is feasible for analyzing risk in non-trivial settings. We look into the feasibility of CIRA for analyzing privacy risks in a case study of an identity management system. Privacy is "too complicated a concept to be boiled down to a single essence" [21]. We agree with

the view of Solove [21] that it is important to understand the socially recognized activities that cause privacy problems to an individual in order to protect it. As the data collected using CIRA will be sensitive and confidential, data is collected through Case Study Role Play (CSRP). CSRP is developed from the integration of case study [26], persona [6] and role play [25]. Personas are "hypothetical archetypes of actual users" and embody their goals [6]. Each role as described in the persona is played by a real person. Using CSRP, data is collected from the individuals playing the role of fictitious characters rather than from an operational setting. In this paper, we have extended the previous work on CIRA by (1) improving the data collection and analysis phase, and (2) showing that it is feasible to use CIRA in non-trivial settings. Our work has contributed to the development of CIRA and helped to identify practical problems that can be addressed in future research.

The rest of the paper is organized as follows. Related work is given in Sect. 2 followed by a description of the case in Sect. 3. In Sect. 4, we present the analysis of the case. We further present and discuss the result of our analysis in Sect. 5. Sect. 6 concludes the paper.

2 Related Work

There are many classical risk management approaches and guidelines. Usually, in these approaches, risk is specified as a combination of likelihood and consequence. The ISO/IEC 27005 [13] standard (its new version ISO/IEC 27005:2011), the ISO 31000 [12] standard (that supersedes AS/NZS 4360:2004 [2]) and NIST 800-39 [16] provide the guidance on the entire risk management process. NIST 800-39 [16] supersedes NIST SP 800-30 [22]; its revised version NIST 800-30 Rev. 1 [17] is a supporting document to NIST 800-39. CORAS [15] is a model based method that uses Unified Modeling Language (UML) for security risk analysis. ISRAM [14] is a survey based model to analyze risk in information security; surveys are conducted for gathering probability and consequence. In Risk IT [10] framework (which is integrated into COBIT 5 [11]), risk is estimated as the combination of frequency (rate by which an event occurs over a given period of time) and magnitude of IT risk scenarios. In RAMCAP [1] (its updated version RAMCAP Plus), risk is estimated as the combination of threat, vulnerability and consequence. Cox has shown the limitations of estimating risk as the combination of threat, vulnerability and consequence [7].

There are several methods that specifically look into privacy risks, and are usually called Privacy Impact Assessment (PIA). For instance, there are Privacy Impact Guidelines of the Treasury Board of Canada Secretariat [18] and PIA of the Information Commissioner's Office, United Kingdom [9]. PIA is a "systematic process for evaluating the potential effects on privacy of a project, initiative, or proposed system or scheme" [24]. It helps to identify and manage privacy risks for an organization that deals with personal data of its stakeholders. However, these methods usually do not attribute the events to people. Wright [24] states

that PIA should be integrated into risk management along with other strategic planning tools.

The CIRA Method [19] identifies stakeholders, their actions and consequences of actions in terms of perceived value changes to the utility factors that characterize the risk situation. The idea being that risk is the combination of the strength of the force that motivates the stakeholder that is in the position to trigger the action to send the risk owner to an undesirable state and the magnitude of this undesirability. Risk magnitude is related to the degree of change to perceived utility caused by potential state changes.

3 Case Description: NorgID Identity Management System

The case description is fictitious but the design of the system is inspired by MinID [8]. The Identity Management System (IdMS) helps to manage the partial identities of end-users. IdMS usually consists of three class of stakeholders: End-user, Identity Provider (IdP) and Service Provider (SP). IdP is the organization that issues the credentials/ electronic identity to the end-user. SP is the organization that provides services to end-user after verifying their identities.

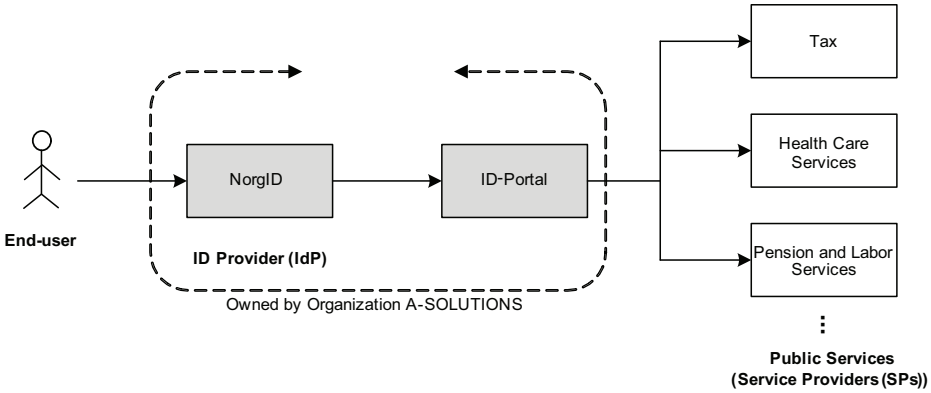

Fig. 1. NorgID Identity Management System.

A-SOLUTIONS is an organization with 20 employees that manages a federated IdMS. It developed an authentication system called NorgID and a portal (ID-Portal). Their goal is to provide secure access to digital public services. NorgID is one of the IdPs which provides authentication for logging on to a federation called 'ID-portal' as shown in Fig. 1. It provides the end-user cross-domain Single Sign-On (SSO), i.e. the end-user needs to authenticate only once and can gain access to many services by using the portal such as tax, health care, pension, labor and other eGovernment services. The end-user can log on

to the ID-portal using NorgID, by providing his personal ID, a password and a one-time PIN code. NorgID uses two databases: (a) for storing personal data about the users and (b) for storing logs containing usage of IdMS for each user (the details regarding the collected information are not mentioned in the privacy policy). The personal information collected includes his social security number, PIN-codes, password, email address, telephone number and address. NorgID has been quickly and widely adopted because of its easy access and features that have convinced enough people to use the application.

4 Analyzing Privacy Risks Using CIRA

In this section, we first provide the assumptions and considerations, along with the scoping for the risk analysis activity. We provide a brief summary of the method along with the steps for data collection (1-9) and analysis (10-13). We then implement the procedure on the given case of an IdMS. The analysis focuses on the risks faced by an end-user.

4.1 Assumptions and Considerations

For investigating the case, we used the Case Study Role Play (CSRP) method. We developed personas of the stakeholders based on empirical data collected for the representative stakeholders. However, for instance, in the case of a hacker, as the empirical data might not be easily elicitable, we used assumption persona [3]. According to Atzeni et al. [3], the assumptions may be derived from different sources of data for the type of individuals that are known to attack the systems. The scenarios were written to provide background information of the role to the participants. We assumed that the participants are honest when interacting with the risk analyst. During the data collection phase, the participants were presented with a set consisting of 3 relevant utility factors. We also asked the participants to provide other factors that they valued or gave them perceived benefit. However, for the simplification of the case we have not considered those factors.

4.2 Scoping

Scoping consists of the activities used to determine the boundary for the risk analysis activity. We (as the risk analyst) assumed that the system is in a certain initial state. Moreover, we focused on privacy risk events that are caused by the intentional behavior of a stakeholder.

4.3 Summary of CIRA

CIRA identifies stakeholders, actions and perceived expected consequences that characterize the risk situation. In CIRA, a stakeholder is an individual (i.e. physical person) that has some interest in the outcome of actions that are taking

place within the scope of significance. There are two classes of stakeholders: the strategy owner and the risk owner. Strategy owner is the stakeholder who is capable of triggering an action to increase his perceived benefit. Typically, each stakeholder has associated a collection of actions that he owns. The risk owner is the stakeholder whose perspective we consider when performing the risk analysis, i.e., he is the stakeholder at risk. By utility, we mean the benefit as perceived by the corresponding stakeholder. Utility comprises of utility factors. Chule et. al. [4] identify the utility factors relevant for our work. Each factor captures a specific aspect of utility e.g. prospect of wealth, reputation, social relationship. The procedure is as given in Table 1 along with the approximate time required for each of the steps when implementing the NorgID case study (the required time will be further explained in Sect. 5).

Table 1. Procedure in CIRA with approximate time required for each step when implementing NorgID IdMS.

Steps	**Time (mins)**
1. Identify the risk owner (includes development of persona)	30
2. Identify the risk owners' key utility factors	30
3. Given an intuition of the scope/ system- identify the kind of strategies/ operations which can potentially influence the above utility factors	30
4. Identify roles/ functions that may have the opportunities and capabilities to perform these operations	60
5. Identify the named strategy owner(s) that can take on this role (includes development of persona)	90
6. Identify the utility factors of interest to this strategy owner(s)	90
7. Determine how the utility factors can be operationalized	240
8. Determine how the utility factors are weighted by each of the stakeholders	120
9. Determine how the various operations result in changes to the utility factors for each of the stakeholders	280
10. Estimate the utility for each stakeholder	20
11. Compute the incentives	15
12. Determine risk	15
13. Evaluate risk	210

4.4 Implementing the CIRA Procedure

The application of CIRA to the NorgID IdMS is presented below.

1. Identify the risk owner. At first we need to determine the risk owner. The user (Bob) is the risk owner. We assume he represents the general users of NorgID. The persona of Bob is given in Table. 2.

2. Identify the risk owners' key utility factors. This step consists of determining the key utility factors for the risk owner.

We presented Bob with three utility factors: privacy, satisfaction from the service and usability along with the explanation for each. We collected his opinion on whether he thought (as a user of NorgID), these factors are important and would give him perceived benefit.

3. Given an intuition of the scope/ system- identify the kind/ classes of operations/ strategies which can potentially influence the above utility factors. For determining the strategies, we look into the taxonomy of activities that cause privacy problems as provided by Solove [21]. The strategies that we considered are:

- Secondary use of Bob's information (**SecUse**): It is related with using Bob's information for another purpose than that is mentioned in the policy without getting his consent.
- Breach of confidentiality of Bob's information: It is "breaking a promise to keep a person's information confidential" [21]. We consider two strategies that can lead to breach of confidentiality: Sharing credentials (**ShareCred**) and Stealing Information (**StealInfo**).

4. Identify the roles/ functions that may have the opportunities and capabilities to perform these operations. There can be many strategy owners capable of executing these strategies. However, for this paper we consider only three stakeholders as the objective is to show the feasibility of the CIRA method. The stakeholders are CEO and System Administrator of A-SOLUTIONS, and a hacker capable of executing SecUse, ShareCred and StealInfo operations respectively.

5. Identify the named strategy owner(s) that can take on this role. In this step, we pin point the strategy owner(s) that are in the position of executing the above strategies. We consider the stakeholders: John (CEO), Nora (System Admin) and X (Hacker). Their personas are provided in Table 2.

6. Identify the utility factors of interest to this strategy owner(s). In CIRA, as we consider the perception of an individual, each relevant stakeholder is an expert. Like before, we provided a list of utility factors for John, Nora and Hacker X to choose from. For the hacker, we identified his utility factors from the existing literature [23]. The identified utility factors for John (CEO): privacy reputation, wealth for business continuity, compliance; for Nora (System Admin): availability, trust, free time and for X (Hacker): wealth, status, ego.

7. Determine how the utility factors can be operationalized. For each identified utility factor, we determine the scale, measurement procedure, semantics of values and explain the underlying assumptions, if any. The brief explantion

Table 2. Personas of risk owner and strategy owners.

Role	Name	Description
End-user	Bob	30 years old, local school teacher, regular user of NorgID with general IT knowledge; aware of some privacy issues mainly due to the media coverage of data breaches (associated with services such as social networking and health care).
CEO	John	50 years old, ensures the overall development and relationship with its stakeholders; has motivation to increase the company's service delivery capacity.
System Admin	Nora	29 years old, known for her friendly behavior and highly trusts her co-workers; ensures both the NorgID and ID-Portal are functioning properly and secure; manages the access permission for internal staff to the server; in her absence, to assure that co-workers get proper system function, she usually lets them access servers and even shares important credentials to the server.
Hacker	X	28 years old, skilled in computing and interested in new challenges; to pursue his interest he left his job a year ago and now completely spends his time by gathering knowledge through first-hand experience; wants to earn money and also build status for himself in the so-called hackers' community.

of the metrics presented in Table 3 and Table 4 are a flavor of the metric we used in the analysis for the stakeholders Bob (User) and John (CEO). It is to be noted that different flavors of the metric exist and can be used according to the context. Due to space constraint, we leave out the details of the metrics for the utility factors of Nora (System Admin) and X (Hacker).

8. Determine how the utility factors are weighted by each of the stakeholders. We asked Bob to rank the utility factors based on its importance. Then, for collecting the weights for the utility factors the following question was asked- "Given that you have assigned a weight of 100 to utility factor #1, how much would you assign to utility factor #2, #3 and so on (on a scale of 0-99)?". Bob ranked and assigned weights of 100, 80, 70 to the utility factors privacy, satisfaction and usability respectively as given in Table 5.

Similarly, the weights of the utility factors according to their ranking for each of the strategy owners were also collected. John (CEO) assigned weights of 100, 80 and 50 to the utility factors compliance, privacy reputation and wealth respectively. Nora (System Admin) assigned weights of 100, 80 and 78 to the utility factors service availability, free time and trust respectively. X (Hacker) assigned weights of 100, 90 and 85 to the utility factors wealth, ego and status respectively.

9. Determine how the various operations result in changes to the utility factors for each of the stakeholders (start with risk owner). We assume the system/ environment to be in a fixed initial state and all the players

Table 3. Metrics for the utility factors of the risk owner Bob (User).

Utility factor	Definition	Measurement Procedure
Privacy(%)	It refers to the extent to which you have control over your personal information. Defined by $$1/(1+N) \quad (1)$$ where N- expected/ projected number of incidents per month.	N is obtained from the analysis of the scenario directly or indirectly caused by the events triggered by various stakeholders [19]. If $N = 0$, the value of privacy is 100%; if $N = 1$, the value of privacy decreases to 50% and so on. That is with increasing number of incidents, the value of privacy decreases.
Satisfaction(%)	It refers to the extent to which you perceive the continuance usage of the portal to access services based on your experience. Model as expectation fulfillment relating to function: service availability, support(reponsiveness (scale: %), effectiveness (scale: %)) and service completeness.	Service availability is the number of interactions with a response time of less than 1 second divided by the total number of interactions. Responsiveness is given as $$1/(1+t) \quad (2)$$ where t is the average time in mins required to 'solve' a problem reported by the user. Effectiveness is the 'extent' to which the problem is solved. Service completeness relates to the number of features that the service actually delivers divided by the number of features that the user could reasonably expect (see [19]).
Usability(%)	It refers to the extent to which a user perceives the ease of interaction with the portal. Model as user's past experience with using the service.	The value can be obtained by doing the survey. A scale of 0 to 100% is used, a value of 0 denotes it takes more than 30 mins to get acquainted with the service; 25% denotes it can be done within 20-30 mins; 50% denotes it takes 10-20 mins; 75% it takes less than 10 mins; 100% denotes it takes less than 5 mins.

Table 4. Metrics for the utility factors of the strategy owner John (CEO).

Utility factor	Definition	Measurement Procedure
Privacy Reputation(%)	It refers to the reputation of the company with respect to privacy incidents (e.g. loss, misuse or breach of personal information). Model as user's expectation relating to future behavior of the company in terms of: experience of others and own experience; both defined by $$1/(1+P) \quad (3)$$ where P is the number of privacy incidents.	P is obtained from the survey. If $P = 0$, the value of reputation is 100%; if $P = 1$, the value decreases to 50% and so on. That is with increasing number of incidents, the value of reputation decreases (see [19]).
Wealth(Million €)	The unit for wealth is currency units. The weight for wealth will then specify how much utility each currency unit will give.	It is obtained from the investigation of the entity by the risk analyst.
Compliance(%)	It refers to the extent to which you think the company would benefit by following the rules and regulations. This demonstrates the willingness of the company to take necessary steps to protect the personal information of its stakeholders. Model as percent of compliance with legislation (e.g. Data Protection Act, EU directive).	At first the risk analyst needs to gather the rules that needs to be followed by the company. A value of 0 means that no rules are followed; 25% means that 1/4 of thoes rules are followed; 50% means that half of those rules are followed; 75% means 3/4 of the rules are followed and 100% means all rules are followed.

Table 5. Utility factors for Bob (User).

Rank	Utility factors	Weights
1	Privacy	100
2	Satisfaction	80
3	Usability	70

are utility optimizing. By utility optimizing, we mean that they are optimizing their behavior relative to the weighted sum of the elements in their utility factor vector. For each of the identified utility factors, we determine the initial and final values after the strategies of the players are executed (for the utility factors' valuation, we utilize the metrics explained above). We use the additive utility function of MAUT to estimate the utility. The additive utility function for a given player is defined to be the weighted average of its individual utility functions [5] given as:

$$U = \sum_{k=1}^{m} w_k \cdot u(a_k) \tag{4}$$

where, m is the number of utility factors of the player, w_k is the assigned weight of utility factor a_k and $\sum_{k=1}^{m} w_k = 1$, and $u(a_k)$ is the utility function for the utility factor 'a_k'.

Table 6. Final Values of the Utility Factors after the Strategy of the Strategy Owners are Executed.

				Final Values		
Stakeholders				**John**	**Nora**	**X-Hacker**
	Utility Factors	**Wts**	**IV**	**SecUse**	**ShareCred**	**StealInfo**
Bob(User)	Privacy(%)	0.40	100	8	17	5
	Satisfaction(%)	0.32	72	74	74	74
	Availability (%)	0.33	85	87	87	87
	Support (%)	0.33	52	55	55	55
	Responsiveness (%)	0.50	14	17	17	17
	Effectiveness (%)	0.50	90	92	92	92
	Service Completeness(%)	0.33	80	82	82	82
	Usability(%)	0.28	80	80	80	80
John(CEO)	Compliance(%)	0.43	80	60		
	Privacy Reputation(%)	0.35	67	15		
	Experience of others(%)	0.50	33	9		
	Own experience(%)	0.50	100	20		
	Wealth(Million €)	0.22	5	25		
Nora(Sys Adm)	Service Availability(%)	0.39	85		87	
	Free time(%)	0.31	0		30	
	Trust(%)	0.30	50		90	
X(Hacker)	Wealth(Thousand €)	0.36	0			50
	Ego(%)	0.33	40			95
	Status (%)	0.31	50			85

For our case study, Table 6 depicts the normalized weights (for the assigned weights in Step 8) for the utility factors, its initial value (IV) and its final values, if the strategies of the stakeholders were to be executed. For the other elements comprising the utility factors, we make the assumption that the stakeholders perceive each of these to be equally important. The values for the metrics are obtained either based on our investigation or by conducting interviews/surveys with the participants. Usually, the individual utility functions (i.e. utility factors in our case) are assigned values in the interval of 0 (worst) to 1 (best) when using MAUT. For instance, in our case, we can easily compress the wealth to the interval 0 to 1. However, this would not be particularly helpful as most of the values will be clustered right at the end. Thus, it is more intuitive to utilize the given scales for the utility factors' valuation. Moreover, the units of the weights are such that the utility is unit less. Next, the values for each of the stakeholders are determined.

For Bob (User). We determine the values of the first two utility factors for Bob from our investigation and the last one (usability) is based on the survey. To determine the value of privacy to the user, we investigated the number of privacy incidents at each state. Our findings are based on several studies on issues such as how secondary usage of data and breach of confidentiality will impact the end-user. Based on our study, $N = 0$ per month at the initial state. $N = 11$, $N = 5$ and $N = 20$ when John, Nora and Hacker X use their respective strategies. By instantiating (1) with the value of N, we obtain the IV of privacy as 100% and its final values as 8%, 17% and 5% respectively.

Note that the values for satisfaction are obtained using the techniques borrowed from MAUT and from our investigation. For support (an element of satisfaction), the values for the responsiveness are obtained after instantiating (2) with $t = 6$ at the initial state and $t = 5$ when the other strategies of the stakeholders are executed. Thus, responsiveness increased from the IV of 14% to 17% for all three strategies. Besides, it was determined that effectiveness also increased from 90% to 92% when the three strategies of the stakeholders are executed. We then evaluate the values for support instantiating (4) with the obtained values of responsiveness and effectiveness: for the IV as 0.50*14+0.50*90 = 52%. Similarly, the final values for the three strategies are evaluated as 55%. The following values were determined for the other elements of satisfaction: availability increases from 85% to 87% and service completeness increases from 80% to 82% after the three strategies are executed. Thus, using (4) and the values determined for the other elements comprising our satisfaction utility factor, the obtained IV is 72% and the final values for the other strategies are evaluated as 74%. The value of usability as obtained from Bob was 80% for all cases.

Due to lack of space, we leave out the details of the computations of changes to the utility factors belonging to the other stakeholders. The results can be found in Table 6.

10. Estimate the utility. We again use the techniques from MAUT to estimate the utility for each of the strategies for each player using (4). We make the

Table 7. Matrix of Utilities and Change in Utilities w.r.t. Strategy of the Strategy Owners.

	Utilities				**Changes in Utilities (Δ)**		
Stakeholders	**IV**	**SecUse**	**ShareCred**	**StealInfo**	**SecUse**	**ShareCred**	**StealInfo**
Bob(User)	85	49	53	48	-36	-32	-37
John(CEO)	59	37			-22		
Nora(Sys Admin)	48		70			22	
X(Hacker)	29			76			47

simplifying assumption that utility is linear. For our case study, we use (4) to compute the utilities for the stakeholders with the values given in Table 6. In the initial state, the utilities are given as follows:
For Bob (User): $0.40 \cdot 100 + 0.32 \cdot 72 + 0.28 \cdot 80 = 85$
For John (CEO): $0.43 \cdot 80 + 0.35 \cdot 67 + 0.22 \cdot 5 = 59$
Similarly, for other stakeholders, the utilities are obtained as given in Table 7.

11. Compute the incentives. We need to compute the incentives (i.e. changes in utilities) for each of the strategies for each player. The change in utility Δ is the difference between the utility of the player in the state resulting from strategy use and the initial state. In our case study, from Table 7, when John uses the SecUse option, Δ for Bob and himself are -36 and -22 respectively. When Nora uses the ShareCred option, the Δ for Bob and herself are -32 and 22 respectively. In addition, when the hacker uses the StealInfo operation, the Δ for Bob and himself are -37 and 47 respectively.

12. Determine risk. This can be achieved by investigating each of the strategies with respect to sign and magnitude of the changes determined in the previous step. In our case study, when John uses the SecUse option, it results in a negative change in utility for both the players (falls in the third quadrant in the incentive graph as shown in Fig. 2). Thus, we know it is an undesirable situation for both the players and they both want to move out of this quadrant. Thus, this might result in co-operation. However, Nora's degree of desirability to play the ShareCred is slightly more as it leads her to a better position with a gain of 22. In this case, 22 is the strength of the force that motivates Nora to send Bob to an undesirable state and -32 is the magnitude of this undesirability and the combination of these is the risk (-32, 22). Similarly, it is clear that the Hacker X's degree of desirability to play the StealInfo is high as it leads him to a better position with a gain of 47 and -37 is the magnitude of the undesirability faced by Bob, which results in the risk (-37, 47).

13. Evaluate risk. We identity the risk acceptance and rejection criteria for the risk owner to determine whether a specified level of risk is acceptable or not. In our model, we make the simplifying assumption that all strategy owners will need the same time to act if they have the same magnitude of incentive.

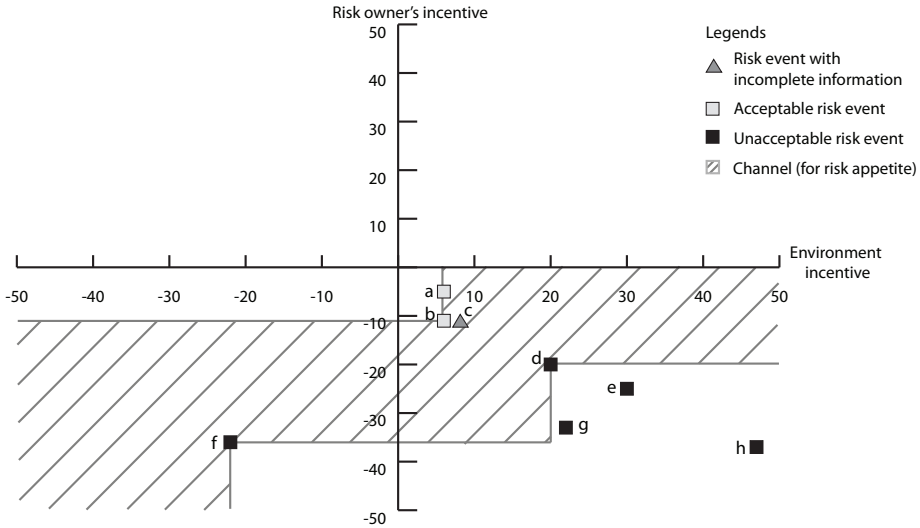

Fig. 2. The Incentive graph

Strategies will be executed in decreasing order of utility as perceived by each of the strategy owners.

We presented Bob with following risk pairs: a. (-5,6), b. (-11, 6), c. (-11, 8), d. (-20,20), e. (-28, 30) along with the ones determined in Step 12, which are f. (-36,-22), g. (-32,22) and h. (-37, 47) obtained when the strategy owners execute the strategies SecUse, ShareCred and StealInfo respectively. The risk pairs are represented by (C, I_i) where C is the consequence for the risk owner and I refers to how strong is strategy owner i's incentive to make the first move or the magnitude of incentive. For instance, for the risk pair 'b', Bob gets the value of C as -11 when the final values for privacy, satisfaction and usability in the execution of any of the strategies would be 95%, 70% and 50% respectively (keeping the weights of the utility factors and their initial values as obtained before). Note that this is one of the several possible combinations that gives Bob the consequence of -11. Nora has an incentive of 6, when the final values for availability, freedom and trust are 90%, 10% and 53% respectively. Similarly for other stakeholders the possible combinations can be determined.

To determine the risk acceptance criteria, we asked Bob (User): 'How strong a temptation is it acceptable to give a strategy owner to execute the strategy, so as to cause him (i.e. Bob) a given loss?'. From the above risk pairs, he accepted the risk pairs a and b (represented by the light gray square) as shown in Fig. 2. However, for the risk pair, c he was willing to accept the risk only if Nora was in the position of executing the strategy (represented by the triangle) and unsure in case other strategy owners executed their strategy. The remaining risk pairs were not acceptable to him (represented by the black square).

5 Results and Discussion

Our findings can be grouped in the following categories: (1) application of CIRA to NorgID IdMS, (2) feasibility of CIRA in terms of its complexity and risk analyst effort required, (3) improvements made and (4) some limitations of CIRA that require further work. Application of CIRA to NorgID IdMS, resulted in the determination of risks faced by the risk owner. We were further able to represent acceptable/ unacceptable risk events by means of an incentive graph which was easy to communicate to the risk owner.

Assuming we have n stakeholders, each stakeholder owns s strategies and has u utility factors that go into the computation of his utility, then the effort of the various tasks can be estimated as follows: The total number of strategies to be considered will be $n * s$. The total number of utility factors to be considered will be $n * u$. However, in practice, it is expected that utility factors will be taken from a limited set. To determine the risk acceptance criteria, it will suffice to ask the risk owner $n * s$ yes/no (i.e. accept/reject) questions. Thus the complexity of CIRA in terms of human effort will be in the order of

$$n * (u + s) \tag{5}$$

By instantiating (5) with the value of $n = 4$, $s = 1$ and $u = 3$ as in the NorgID case study, we obtain the estimate of complexity as 16. Furthermore, the effort in terms of total amount of time spent in doing the case study was determined to be approximately 27 hrs (which includes the time given in Table 1 along with the time for initial preparation (1 hr), scenario construction to provide the background information of the role to the participants (2 hrs), role play selection and guidance (2 hrs) and documentation (1.5 hrs)). The given hours are approximate values; the values were jotted down only after the actual process was completed. It is clear that steps for determining the changes to the utility factors with respect to the operations (Step 9) and the operationalization of utility factors (Step 7) required the highest amount of time i.e. approximately 280 and 240 mins respectively. When the problem space grows, for instance the values of $n = 8$, $s = 10$ and $u = 5$, we would expect that the risk analyst would have to spend in the order of 200 hours to complete the analysis. Note that the elapsed time may be longer. CIRA is still in development phase and the steps will be optimized. For e.g. a comprehensive library of utility factors will be developed. It is expected that this library will speed up the data collection phase. Moreover, tools will be developed to support the risk analyst.

Learning from the case study, we discovered the following issues that resulted in improvements: the procedure was updated to ease the data collection process and the data collection manual was developed for the risk analyst. Interviews/ survey responses indicated that it was essential that the risk analyst and the participants have the same understanding of the concepts (e.g. utility factors) used during the data collection phase. Thus, even though a lot of resources were required for instance, in the operationalization of the metrics for the utility factors and also determining their value, we focused on these key issues in order to improve data quality.

The following limitations of CIRA were identified: (1) We have assumed that all the participants are honest when interacting with the risk analyst. However, the fact that they might be reluctant to provide information or give wrong information during the interview/ survey needs further investigation. (2) As metrics have always been a challenge in information security, for some of the utility factors it was difficult to formulate the metrics. Hence, we need to collect definitions of utility factors and perform their validation. (3) To determine whether an obtained set of utility factors represents the complete set for a particular stakeholder in a given context requires further work. (4) More work is also needed in capturing the uncertainties in relation to estimates using interval arithmetic or bounded probabilities instead of point values. (5) When assigning weights, the same scale is used for all the stakeholders. The mapping of scale of one stakeholder with another also needs further investigation. (6) Finally, CIRA tool support.

6 Conclusion

In this paper, we have explored the feasibility of CIRA to analyze risk in a non-trivial setting. The CIRA method is still at an early stage of development. However, the results from our case study suggests that it is possible to use CIRA in such settings, and that the method helps the analyst to get a better understanding of the risks. Our work has contributed to the development of CIRA and helped to identify practical problems that can be addressed in future research.

Acknowledgement. The work reported in this paper is part of the PETweb II project sponsored by The Research Council of Norway under grant 193030/S10. We would like to thank the anonymous reviewers for their valuable comments.

References

[1] ASME Innovative Technologies Institute (ASME-ITI). *RAMCAP(Risk Analysis and Management for Critical Asset Protection) Framework*, May 2006. Version 2.0.

[2] AS/NZS 4360. *Risk management.* AS/NZS, 2004.

[3] Andrea Atzeni, Cesare Cameroni, Shamal Faily, John Lyle, and Ivan Flechais. Here's Johnny: a Methodology for Developing Attacker Personas. *ARES*, pages 722–727, 2011.

[4] Ada S. Chulef, Stephen J. Read, and David A. Walsh. A Hierarchical Taxonomy of Human Goals. *Motivation and Emotion*, 25(3):191–232, 2001.

[5] Robert T. Clemen. *Making Hard Decision: An Introduction to Decision Analysis.* Duxbury, 2nd edition, 1996.

[6] Alan Cooper. *The Inmates Are Running the Asylum.* Macmillan Publishing Co., Inc., Indianapolis, IN, USA, 1999.

[7] Jr. LA Cox. Some limitations of "Risk = Threat x Vulnerability x Consequence" for risk analysis of terrorist attacks. *Risk Analysis*, 28(6):1749–61, 2008.

[8] Difi (Direktoratet for forvaltning og IKT). MinID. http://minid.difi.no/minid/minid.php?lang=en. [Online accessed: 11-2012].
[9] Information Commissioner's Office (ICO). Privacy Impact Assessment Handbook, 2009. Version 2.0.
[10] ISACA, Rolling Meadows. *The Risk IT Framework*, 2009.
[11] ISACA. *COBIT 5: A Business Framework for the Governance and Management of Enterprise IT*. IT Governance Institute, 2012.
[12] ISO 31000. *Risk Management – Principles and Guidelines*, 2009.
[13] ISO/IEC 27005. *Information technology -Security techniques -Information security risk management*. ISO/IEC, 1st edition, 2008.
[14] Bilge Karabacak and Ibrahim Sogukpinar. ISRAM: information security risk analysis method. *Computers & Security*, 24(2):147 – 159, 2005.
[15] Mass Soldal Lund, Bjørnar Solhaug, and Ketil Stølen. A Guided Tour of the CORAS Method. In *Model-Driven Risk Analysis*, pages 23–43. Springer Berlin Heidelberg, 2011.
[16] NIST. *NIST SP 800-39, Managing Information Security Risk - Organization, Mission, and Information System View*, 2011.
[17] NIST and U.S. Department of Commerce. *NIST SP 800-30 Revision 1, Guide for Conducting Risk Assessments*, September 2012.
[18] Treasury Board of Canada Secretariat. Privacy Impact Assessment Guidelines: A Framework to Manage Privacy Risks Guidelines. http://www.tbs-sct.gc.ca, April 2012. [Online accessed: 1-2013].
[19] Lisa Rajbhandari and Einar Snekkenes. Intended Actions: Risk Is Conflicting Incentives. In Dieter Gollmann and Felix Freiling, editors, *Information Security*, volume 7483 of *Lecture Notes in Computer Science*, pages 370–386. Springer Berlin / Heidelberg, 2012.
[20] James Shanteau and Thomas R. Stewart. Why study expert decision making? Some historical perspectives and comments. *Organizational Behavior and Human Decision Processes*, 53(2):95–106, Nov 1992.
[21] Danile J. Solove. A Taxonomy of Privacy. *University of Pennsylvania Law Review*, 154(3):477, January 2006. GWU Law School Public Law Research Paper No. 129.
[22] Gary Stoneburner, Alice Goguen, and Alexis Feringa. *NIST SP 800-30, Risk Management Guide for Information Technology*. NIST, July 2002.
[23] The Honeynet Project. *Know Your Enemy*. Addison-Wesley, 2nd edition, 2004.
[24] David Wright. Should privacy impact assessments be mandatory? *Commun. ACM*, 54(8):121–131, August 2011.
[25] Krysia M Yardley-Matwiejczuk. *Role play: theory and practice*. Sage Publications Limited, 1997.
[26] Robert K. Yin. *Case Study Research: Design and Methods*, volume 5 of *Applied Social Research Method Series*. Sage, 4th edition, 2009.