Secalyser – A System to Plan Training for Employees

J. Schlüter and S. Teufel

University of Fribourg/Switzerland
e-mail: jan.schlueter@unifr.ch, stephanie.teufel@unifr.ch

Abstract

Secalyser uses an approach to measure knowledge and behaviour of employees. It takes into consideration the information technology, especially in with regard to human security aspects. Furthermore secalyser is able to determine the need of action recommending actions for automatically selected groups to improve information security. The evaluation of this need of action is done by one of two specially developed survey tools, as well as a new interpolation system integrated into secalyser.

Keywords

Measure human security, plan training, secalyser, improve information security, identification of the need of action

1. Introduction

In most cases the influence of employees' security knowledge and their behaviour is underestimated by companies. This is due to the fact that it is much easier to change some parts of the technical environment, or replace a software package with a newer one, than recognising that the employees are the weakest part in the information security chain (Künzler, 2002; Wylder, 2004).

For this reason, the security level can only be as high as its weakest parts, thus it is important to improve the weak parts. In common cases, only a very small number of employees allegorise a huge part of the company's security risk. This is caused by either a very special position of some employees (e.g. members of the business management or the information systems administration), or large substandard security knowledge. The goals are to find out this small number of persons, to increase the security level with less company resources, and to select individual training for each employee of this group (Zürcher Tagung, 2007; Deswarte et al., 2004).

In huge companies and concerns, it is very hard to plan different trainings in a short period of time. This becomes more and more complicated when the training is not only for a single or a small group of persons, but for a larger group. If many equivalent training courses take place parallel, it is difficult to optimise the courses for each person while keeping the training size small. This is due to already existing meetings, as well as employees duties. These circumstances will also be accounted for in secalyser's calculation (Wylie and Grothe, 1996).

One of the biggest problems is to bring terms like Security Culture and Security Awareness to the companies. A frequently noticed behaviour is that companies start to think about security only after having their first incidents, and stop their actions a short time later. It is essential to develop new approaches to make sure that companies are able to handle security as a process, as well as to understand how to enhance the employees' Security Awareness. Security Culture in usual companies is almost unknown and is often examined as unimportant by the management (Schlienger, 2006).

2. Causations of human security risks

To understand which part of the human security risk can be enhanced by using secalyser, it is important to categorise the single causes of human security risks.

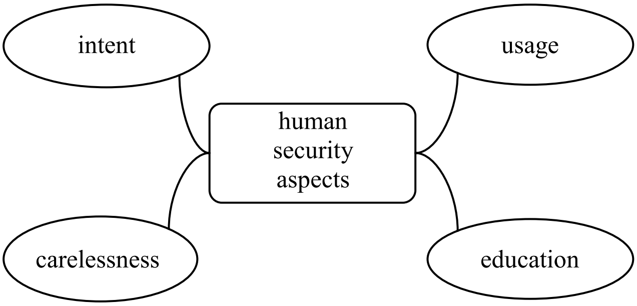

Figure 1: Human Security Aspects

2.1. Intent

An employee, who is an intentional security risk, is the most dangerous employee a company can have. An employee in this group has the knowledge to do right, but intentionally wants to do damage to the company. The only possibility to stop these employees being a security risk is to identify them and to dismiss them. Due to the fact that these persons have the knowledge to pass a test, and there is no common reason for their behaviour, it is not possible notice them automatically (Berndsen, 1997; Parker, 1998).

2.2. Usage

Operational procedures and processes inside a company may change from time to time, and sometimes a reason for such a change is a security issue. Employees who do not recognise this, or who think that the previous procedure is better and do not see the necessity to change their behaviour, need at least an advice or a better training. This teaches these persons the reason for the changed procedure. Normally these persons (who often have worked there for a longer time) would change their behaviour in a short time (Kaufmann et al., 2007).

At present, secalyser does not support finding these employees, but there is a new idea to handle this case with the help of the Business Process Modelling Notation (BPMN), but this would only work with big companies who have all their processes written down in BPMN (Schlüter, 2007).

2.3. Carelessness

In comparison to 2.2, the employees' profiles are the same, but they know about the change and they know exactly why the company changed the process, however "most human beings enjoy the experience of free will" (Reason, 1990, p. 235). This group of employees does not change their behaviour, because they do not care about it and they want to keep their work as simple as it was. In some cases it is hard do differentiate between employees in this group and employees from 2.2. In special cases, in particular when the employees are informed about the new process a couple of times, the employees can be classified in group 2.1 too (Künzler, 2002).

By reason that this group is a mixture of a separate group (when they change their behaviour after being advised), and an intersection of the other groups, it is hard to identify these groups of employees automatically. To identify this group it would be necessary to log the whole communication between the employees in a digital way, and to analyse the data. This function could be implemented through a Customer-Relationship-Management-like system (CRM), but set-up for employees and not for customers. Currently this function is not available, but it is planned for future development.

2.4. Education

This group is well-defined from the other ones by secalyser. By using a special survey and interpolation system, and given access to the company's master data, secalyser is able to find out about the group of employees who need training, and suggests courses for each person. It is also possible to schedule the best time for training if there is a digital calendar for each employee available.

The way secalyser does this analysis and scheduling is described in the next chapters.

3. Survey

The survey is used to test the current security level of an employee. To make it possible for secalyser to get automatic access to the survey results, the survey has to be digital. It would be possible to source out the survey system to an Application Service Provider (ASP), or to host the survey system in-house. For the second possibility there are two survey systems developed within the secalyser project; the first for smaller companies (PHP Hypertext Processor[1] written and without Database Management System), and the second for bigger ones (Java Platform Enterprise Edition[2] based with Java Database Connectivity[3]).

Every employee who should be benchmarked by the system has to participate in at least one survey. The ideal case would be that every employee who has access to computers at work has participated in more than three surveys, where the last survey should not be older than about a year. Due to the fact that we need at least two years to evaluate the change of an employee's security level, we are unable to give exact values here.

The survey is divided into two different parts. The general approach of this kind of survey was already tested in Schlienger (2006).

3.1. Common part

In this non-technical part, only common questions are asked. This part gives an overview of the general security affinity of an employee, without knowing his or her information technology skills. To get reasonable results, this common part has to be adapted for different branches and indirect guidelines of the benchmarked company. An example would be to ask for the knowledge of the company's security policy, even if the company does not have one.

At present we are only able to check the plausibility of the employees' answers by generally known methods such as cross questions or asking some facts with other formulations in hope that the employee does not recognise the interrelation. However, it is impossible to assure that the employees' answers agree with their real opinion. The intricacy is even more increased there is a big difference between the cognitive knowledge and the real behaviour. This means, for example, that everyone knows that he has to stop at red traffic lights before crossing the street, but many people behave differently.

By defining aligned calculation instructions for the adapted survey, we get an evaluation key to interpret this first part. While a point based result, which can be added to the overall score, would not represent the idea of the general plan described above, we need a method which has an effect on the whole survey, but does not improve the score, when there is no technical background. Because of this, in difference to the second and more technical part, the score is percentage based and the different answers are evaluated in relation to each other. The percentage result of

[1] PHP is a scripting language for creating dynamic websites

[2] Server platform for Java programming language

[3] JDBC is a library to connect to Database Management Systems in Java

this first part can be used to be applied on the point based score in the second part. The advantage is that bad results in the second part are not enhanced by a good general behaviour, but a bad general behaviour degrades the achieved points. This approach tries to consider that the employees' behaviour in new situations, which are not trained and which are not asked for in the survey, will be dissolved better when the general security relevant behaviour of the employee is better.

In our tests we assume that a maximum loss of 40 percent, that means a scope from 60 to 100 percent, is an expedient clue for the adaption of the evaluation key. While adapting the key, it should be estimated that in most cases the function is continuously increasing with a right-hand limit.

3.2. Technical part

The technical part concerns critical situations in detail. For every correct answered question the employee gets points. Each question has a different number of answers and each answer can give a definable number of points. Each questions aims to duplicate a special situation which can occur in real life. This "Simulator learning" approach is one of the best ways to make the trainees learn from their errors (Reason, 1990, p. 244-246). Because of this, it is very important to correct them afterward. Until now this has to be done manually.

Due to the fact, that some questions do not make sense in some companies, the survey has to be adapted. However, in comparison to the common part not the whole question catalogue has to be rewritten, but only some nonsensical questions have to be removed.

4. Implementation and program flow

Secaliser is designed as shown in Figure 2. The two central elements are the *Database Management System (DBMS),* and the *Reasoner*. The *DBMS* is used as a database warehouse to archive actual and historical data, making it possible to set all data into relation. Due to this fact, the Start-Schema, which was developed for this purpose, is used. The *Reasoner* contains the whole logic, catches the data from the *DBMS* and the *WebDAV Interpreter* and sends the data to the RPT-Report engine to generate the *Report*. Due to this fact, it is not necessary to store historical data from the *Microsoft Exchange Server*, this system is not connected with the *DBMS*.

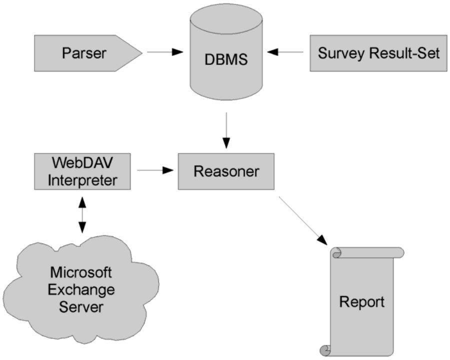

Figure 2: Implementation Design

Secalyser has four completely independent functions which can each be started manually. Additionally there is a start function which initiates each of these four functions in series, but in most cases it would be useful just to start the function which is actually needed, due of the computing time of each function.

4.1. Synchronise ERP-Data

Secalyser needs a copy of the actual ERP Master Data. This is done by automatically importing an XML-Dump of the ERP System.

The format of the dump can be adapted by changing the style sheet definition. In our example we built style sheets named "*company*" and "*employee*". At present the *company* style sheet only contains definitions for the types named *nonEmptyString* (a string with at least one character) and *percent* (a double value from 0 to 100).

Due to the fact that we are currently enhancing secalyser's features, we will need to link further style sheets with the company – the XSD Design will be more comprehensible after this has been done.

The *employee* definition contains with the four necessary values:

- a unique identifier (in this example; email)
- the employee's date of birth
- the employee's skill-level
- the employee's importance level

While the first two values do not need any explanation, the other two are described in sections 4.1.1 and 4.1.2.

4.1.1. Skill

Every employee has a job position, which makes it possible to declare indirectly a skill area where each employee is in a certain position. In combination with the employee's age, a conclusion of the employee's career is possible.

The following example should exemplify this theory:

- A employee has a job position as "consultant"
- The job position as consultant has a skill level of 70%
- The employee is 35 years old

Due to the fact that the employee is quite young and that the job as consultant gives the possibility for promotion, the skill level can be adjusted upwards. This level of adjustment is calculated by the age of the employee and the possibility of promotion.

Of course this is part of a first approximation and is only used if there are not enough measured results from at least two former surveys.

4.1.2. Importance

The importance factor describes how important the information technology is for a certain job position. For example, the company's CIO would have a higher importance factor than the concierge. Indirectly this factor gives information about the necessity of training for certain employees.

4.2. Analyse the survey

The results of the used survey system are read from a flat-file or database system and the score is written to the database system. In this step there is no interpolation, because all data being changed comes from the present survey.

As described in chapter 3, each survey is, at least in some points, different from the other ones, but the way on how to evaluate the results is identical, however there are some specifics to keep in mind:

- A bigger percentaged scope in the common part will automatically induce lower scores of the average employee.
- Unbalanced score distribution indicates either a number of groups of employees that you cannot compare and should look at separately, or a few non expedient questions in the technical part.

4.3. Interpolate the results

All employees' scores which are not up to date (surveyed in the current month) are interpolated to the current month. To get good results in interpolating, one of two different options is used. Which option is used depends on the amount of data that is available from former surveys.

Keep in mind that all employees who never took part in a survey cannot be included in the whole process. There is at least one survey necessary.

4.3.1. None, or too little data available

In this case, which is normally used in the time of launching the tool at a company (if there is no data entered manually), the gradient score level is calculated by using statistical average values selected by the known values of "age" and "skill".

Of course this case is the worst case for the accuracy of data, but without knowing further details no exact calculation is possible. Nevertheless this approach shows a way to handle even this almost unknown part of employees.

4.3.2. Sufficient data available

If there is sufficient data available (the amount of data assessed as sufficient depends on secalyser's configuration), the former learning curve of the employee is analysed, and the interpolation will be based on the former curve behaviour.

Secalyser has the ability to differentiate automatically between the following mathematical approximations and uses the best fitting one depending on the standard deviation of each approximation with the real curve.

The following modes of approximation are known by secalyser:

- linear
- potential
- polynomial (fifth grade Taylor polynomial)
- trend (uses a linear approximation of the last three scores)

Each mode of interpolation can be disabled manually.

4.4. Plan trainings and cluster the employees

After the current score is calculated in the last step, the employees have to be clustered into groups and the different training plans have to be assigned to those groups – depending on the necessity.

The problem is to find a way to combine the following three approaches:

4.4.1. Optimise by date

This optimisation uses the private calendar of each employee (a MAPI interface for connecting to Microsoft Exchange servers is part of secalyser), and builds an index which describes how good a certain date fits into the calendar of each person.

It also considers that it is more secure to let the trainings take place as soon as possible, and additionally, it can be set that persons with a higher job position are weighted more.

The problem is the base period, which has to be limited to get results in an acceptable time, and the fact that the calculation is NP-Complete.

NP-Complete in this context means that each group has to be checked against each other group. This results in a non-polynomial calculating time – depending on the number of employees, the maximum and minimum size of each group, and the number of trainings.

4.4.2. Optimise by job position

It is assumed that employees in higher job positions would like to have their trainings with other employees in similar job positions. With employees in lower job positions, a similar relation is expected.

Due to this assumption the groups can be organised by their job position to make the training more efficient.

4.4.3. Optimise by training concentration

It is assumed that trainings are more efficient if there are no free days within training. Because of this assumption, secalyser has a calendar with all public holidays to calculate the number of free days within training.

4.5. Optimisation

To combine the different approaches in 4.4.1, 4.4.2 and 4.4.3, secalyser can disable each approach and define the order that every algorithm is used in.

It is recommended not to use 4.4.1 as the first algorithm while calculating for 10.000 employees and more (because of the NP-Complete behaviour). After using one of the other algorithms, the number of scenarios shrinks so much that it should be no problem to use the optimisation by date with even more than 30.000 employees, although the calculation process may last some moments.

5. Conclusion

The system secalyser has the ability to measure and interpolate employee's information technology and security relevant knowledge, and to plan trainings for them.

An optimisation is in this respect useful, because not every employee will need training as immediately as other employees do. Not only does the trainings cause high costs, but there is also the internal personal costs caused by absence. Secalyser will reduce this overhead and bring the best fitting training to each employee.

The goal is to improve or at least maintain the security level, but to reduce costs by avoiding unnecessary trainings.

References

Berndsen, D. (1997), "Sabotage. Die bewusste und absichtliche Schädigung von Organisationen durch ihre Mitarbeiter", Peter Lang Verlagsgruppe, ISBN 978-3-631-31738-9.

Deswarte, Y., Cuppens, F., Jajodia, J. and Wang, L. (2004). "Security and Protection in Information Processing Systems", Kluwer Academic Publishers, ISBN 978-1402081422.

Kaufmann, F., Akermann, L. and Schmid, O. (2007), "Sicherheitsrisiko Mensch", Gruppenarbeit, unpublished.

Künzler, C. (2002). "Kompetenzförderliche Sicherheitskultur. Ganzheitliche Gestaltung risikoreicher Arbeitssysteme", Mensch – Technik – Organisation, 1. Auflage, vdf Hochschulverlag an der ETH Zürich, ISBN 978-3728128577.

Koen, M. (2005). "The Human Factor", unpublished.

Parker, D. B. (1998). "Fighting Computer Crime. A New Framework For Protecting Information", John Wiley and Sons, ISBN 978-0-471-16378-7.

Reason, J. (1990). "Human Error", Cambridge University Press, 1st edition, ISBN 978-0521314190.

Schlienger, T. (2006). "Informationssicherheitskultur in Theorie und Praxis", Dissertation, iimt University Press, ISBN 978-3-906428-89-5.

Schlüter, J. (2007). "Automatische Planung von Schulungsmassnahmen zur Minimierung humaner Sicherheitsprobleme im betrieblichen Umfeld", Diplomarbeit, Carl von Ossietzky Universität Oldenburg.

Wylder, J. (2004). “Strategic Information Security”, Auerbach Publications, ISBN 978-0849320415.

Wylie, P. and Grothe, M. (1996). “Problem Employees. How To Improve Their Behaviour and Their Performance”, Piaktus, ISBN 978-0749913670.

Zürcher Tagung, Information Security Society Switzerland (2007). “Firewall Mensch”, http://www.zuerchertagung.ch, panel discussion, 23.9.2007.