DSAPE

Dynamic Security Awareness Program Evaluation

Charalampos Manifavas[1], Konstantinos Fysarakis[2], Konstantinos Rantos[3], and George Hatzivasilis[2]

[1]Dept. of Informatics Engineering, Technological Educational Institute of Crete, Heraklion, Crete, Greece
harryman@ie.teicrete.gr
[2]Dept. of Electronic & Computer Engineering, Technical University of Crete, Chania, Crete, Greece
{kfysarakis, gchatzivasilis}@isc.tuc.gr
[3]Dept. of Computer & Informatics Engineering, Eastern Macedonia and Thrace Institute of Technology, Kavala, Greece
krantos@teikav.edu.gr

Abstract. This paper addresses the importance of continuously evaluating an organization's awareness program and provides guidelines that will help organizations assess their efforts, extending the authors' work in [1]. The proposed methodology evaluates an awareness program considering the most common and essential methods used for delivering awareness material. Key awareness-related processes and accompanying quantitative metrics are identified, along with a methodology for dynamically evaluating the metrics and the overall awareness program as a whole. A software tool is developed, to facilitate the deployment and maintenance of the assessment methods and to formalize their aggregation and evaluation. An organization's security awareness posture is modelled as a dynamic system and the awareness level is calculated and monitored through time via Event Calculus. Furthermore, the tool can be deployed in a multi-agent form, to enable its use by organizations operating through remote offices and distributed locations.

Keywords: security awareness, evaluation methodology, security management, event calculus, JESS, JADE, multi-agent.

1 Introduction

In the context of an enterprise environment, security awareness refers to the knowledge and attitude employees possess regarding the protection of the physical and information assets of their organization. Security awareness is a vital element to the orderly and uninterrupted operation of an organization. Even the most efficient security mechanisms have little value in an organization with no security culture, as the human factor often proves to be the weakest link; though, surprisingly, the importance of appropriate awareness and training is often overlooked [2]. Moreover, the

`

importance of security awareness is bound to increase with the introduction of smart office environments through the deployment of various embedded computing systems. Employees are already insufficiently educated on the risks introduced by new working behaviors (e.g. working in public spaces and/or involving life-logging applications [3]), as technological advancements have outpaced awareness efforts [4]. This "awareness gap" is bound to be exacerbated as we move towards the Internet of Thing (IoT).

At any rate, awareness efforts can be of limited effectiveness unless a needs assessment is conducted prior to deployment, in order to facilitating tailoring the program to the specific organization [5][6]. Moreover, the maturity of the program can play a significant role in its effectiveness; the latter cannot be guaranteed during the first years of deployment. Evaluating the overall information security program of an organization is not enough, as it can only give some indications on the efficacy of its awareness methods; a methodology that focuses specifically on the awareness campaign can provide more detailed and accurate results. Measuring the impact of the awareness campaign is, therefore, vital for ensuring program improvement and continuation through management support, as well as for assessing the awareness team's efforts, providing valuable feedback regarding the effectiveness of the chosen strategy and methods.

There are two factors that have to be considered, in order to assess the effectiveness of a security awareness campaign:

- *Has the information reached the target?* While certain methods provide assurance that the information is bound to reach the target, others rely on the deployment strategy. E.g. asking a person to hand awareness brochures personally to each employee is certainly more effective than leaving them on a desk and asking employees to collect them. On the other hand, awareness material distributed via emails is bound to reach the target.
- *Has the information touched the target?* This is the most important aspect in awareness program evaluation as it assesses how many people actually absorbed the delivered information and, therefore, whether the main aim of the program, which is to create security aware and conscious people, has been achieved.

The above distinction is important and widely cited in related security awareness work, as in [7], where the authors apply a security awareness prototype on an international gold mining company with 25 operations in 11 countries. They use metrics to measure three dimensions of awareness: knowledge (what you know), attitude (what you think) and behavior (what you do). It was identified that the different evaluated factors cannot contribute the same to the final awareness level, thus weights were applied, a technique that is utilized in the work presented here as well.

This paper is structured as follows: Section 2 presents our awareness evaluation framework, along with all the identified metrics, parameters and processes which constitute the "Dynamic Security Awareness Program Evaluation (DSAPE)" methodology, Section 3 describes the accompanying tool, including implementation details and a demonstration of its operation and, finally, the work is concluded in Section 4.

2 The evaluation methodology

The evaluation of a program could be based on qualitative or quantitative techniques or a combination of the above.

Qualitative techniques are mainly used to capture employees' sensation regarding awareness and whether they truly exercise security awareness. Although the interpretation of the results obtained by these techniques can sometimes be subjective and might lead to speculations and conjectures, their significance should not be underestimated. Commonly deployed qualitative techniques include users' feedback, independent observations and silent monitoring of employees' reactions (e.g. during an awareness session).

Quantitative techniques attempt to present the evaluation results in a more objective way and provide benchmarks for future evaluations. Methods that can be deployed are metrics, namely key performance indicators (KPI), which can give a clearer view regarding the effectiveness of a program. However, there are neither standardized, universally accepted and validated methods nor exact figures in the industry that can classify a program as successful or not. What is more, defining quantitative metrics appears to be very difficult for most organizations [5]. This is not surprising since these metrics often involve simplifying a complex socio-technical situation down to numbers or partial order [8].

The evaluation methodology presented in this paper will focus on quantitative techniques; quantifiable and repeatable results are an important factor to consider choosing an effective and useful set of metrics for any relevant evaluation, as indicated by all relevant guidelines (e.g. [9]).

2.1 Evaluation metrics

In the following section, we list some recommended quantitative metrics that will be used in the evaluation methodology, organized in 12 categories. Details on the definition, deployment and marking scheme of the individual metrics can be found in [1].

General metrics.

(a) **Surveys**. Questionnaire-based surveys conducted on technical and security policy issues are one of the most reliable means of measuring a program's effectiveness.
   (i) **M1:** Statistical analysis of monthly surveys on specific organization's divisions
   (ii) **M2:** Statistical analysis of annual surveys

(b) **Awareness/Security Days**: Security days offer a unique opportunity for the awareness team to directly communicate with employees and get their feedback.
   (i) **M3:** Statistical analysis of security days attendance

(c) **Independent observations**. Independent observations on the security behaviour of employees are an important indicator of whether the awareness campaign has touched the target audience.

`

(i) **M4:** Statistical analysis of unsuccessful mock phishing attacks
(ii) **M5:** Statistical analysis of new threat bulletins' readership

(d) **Audit department reports**. Auditing can be used to determine if security awareness related incidents identified by audits are declining. Note that this figure should not include issues that fall within specific roles responsibilities and require training and education, as opposed to awareness [6][10].
   (i) **M6:** Number of security issues related to employees security behavior identified by the audit department

(e) **Risk department reports**. Input from the risk department can be used to identify risks related to security awareness. Risks identified during previous risk assessments should be reduced throughout time.
   (i) **M7:** Number of security issues related to employees security behaviour identified by the risk department.

(f) **Security incidents**. Security incidents are a valid point of reference regarding awareness program evaluation, and their processing should go beyond a simple check on the volume of incidents.
   (i) **M8:** Number of employees who are the source of at least one security incident that stems from non-secure behavior (out of the total number of employees).
   (ii) **M9:** Number of employees who are the source of at least one security incident that falls within their responsibilities but were not identified by them (out of the total number of employees).

Individual module metrics.

(g) **Awareness sessions (workshops)**. This is considered one of the easiest methods to evaluate given the existence of multiple communication paths for getting the required feedback.
   (i) **M10:** Statistical analysis of sessions attendance
   (ii) **M11:** Statistical analysis of sessions effectiveness

(h) **Information security website**. The number of employees who visit the website where information security related content is posted demonstrates users' interest in the corresponding topics.
   (i) **M12:** Statistical analysis of information security website visits

(i) **e-Learning**. Statistics can provide useful information regarding the number of employees visiting, registering, and completing the e-learning program.
   (i) **M13:** Statistical analysis of e-learning program visits
   (ii) **M14:** Statistical analysis of e-learning program registrations
   (iii) **M15:** Statistical analysis of completions

(j) **Emails**. Awareness content delivered through emails is bound to reach the target, but the email may be ignored. A simple technique can be used to measure the method's effectiveness: the content can be structured in such a way so that a link is provided as a follow-up for more information regarding the addressed subject and which can be used to measure readers' interest.
   (i) **M16:** Statistical analysis of email views

(k) **iNotices**. As with emails where content is delivered electronically, links can be provided in iNotices for follow up information.
   (i) **M17:** Statistical analysis of iNotices readings
(l) **Posters**. Measuring posters contribution to the awareness should involve independent observations, combined with electronic means, e.g. the use of QR codes that provide links to additional resources or the URL where the same poster can be found in electronic form so that employees can download it.
   (i) **M18:** Statistical analysis of poster downloads

2.2 Other factors

Weighting.
Weighting of the metrics and their individual categories is also incorporated in the scheme, so that the system can be tailored to each organization's specific needs and environment. We introduce some sample weight values for demonstrative purposes, but these should be appropriately distributed by higher management in cooperation with the awareness team prior to the initial evaluation. Some guidelines are also included, by giving emphasis on parameters pertaining to assessing the organization's security culture. The latter is the most important aspect in awareness program evaluation as it assesses how many people actually absorbed the delivered information and, therefore, whether the main aim of the program, which is to create security aware and conscious people, has been achieved. Moreover, whatever the exact weights decided upon initial evaluation, further fine tuning is to be expected and, in fact, necessary to optimize the accuracy and efficiency of the evaluation method as the program progresses and new iterations are deployed.

Cost.
The proposed framework also considers the cost of implementing and running the various awareness-related mechanisms, to facilitate various types of analyses that will help an organization better evaluate the cost-benefit relationships and other aspects of said mechanisms. This facilitates the comparison of T&A initiatives (e.g. one initiative costing $X and focusing on a subset of awareness mechanisms vs. other initiative costing $Y and focusing on another subset of the mechanisms) and provides valuable information to the decision-making process regarding future directions of the awareness program.

2.3 Evaluation lifecycle

In order to implement a continuous awareness evaluation program, the processes detailed above need to be executed in a structured and timely manner. This is depicted in the evaluation lifecycle below:

1. Personalize the framework (set weights, identify pertinent metrics etc.)
2. Define the baseline (first run of the evaluation)
3. Set goals and milestones

`

4. [Optional] Introduce changes and justifications (e.g. new delivery methods/campaigns and pertinent metrics, abandon failed methods)
5. Monitor
6. Re-evaluate (upon milestones) & assess results
7. Repeat from step 3.

3 The evaluation tool

A tool is provided for a formally validated aggregation of the individual awareness-related processes' evaluation, through their respective metrics, and their cost, in order to produce an overall score. This is accomplished via a model-based framework for dynamic metrics composition and awareness evaluation. In specific, Event Calculus (EC) [11] is applied for modeling the behavior of a dynamic system and calculating its awareness level through time. The resulting overall score is usable both as a benchmark for future iterations of the evaluation program as well as a figure presentable to higher management. Other features include recommendations based on a metric's record, both in terms of absolute value as well as in terms of the value's change over time. Areas with very poor cost-benefit performance are highlighted, including suggestions about specific changes that could help identify and address the causes behind a mechanism's subpar performance (e.g. "Consider revising T&A session material").

In this section, we describe the implementation details of the DSAPE application, the evaluation process of a security awareness program and the recommendation process. We present how our tool can be utilized by higher management in enterprises and demonstrate a use-case with all the aforementioned metrics.

3.1 Implementation Details

The DSAPE tool implementation is based on Event Calculus (EC) [11]. EC is a logic language for representing and reasoning about actions and their effects. Discrete Event Calculus Knowledge Theory (DECKT) [12] is an implementation of EC with the rule engine Jess [13]. DECKT can perform, among others, automated epistemic, temporal and casual reasoning for dynamic domains. DECKT is extended in [14] with real time events, preferences and priorities. The extended DECKT is transformed to an agent's reasoning behavior with a GUI, which is applied to Java Agent Development framework (JADE) [15]. Different security awareness agents communicate with the standardized Agent Communication Language (ACL) [16].

We model the security awareness program along with its modules and metrics as fluents, and the evaluation of metrics as events of EC and implement them in the extended DECKT. Every program, module and metric contains one method in Java that implements the formulas for evaluating its security awareness level and cost (based on the methodology detailed in Section 2). Moreover, we implement rules in Jess that trigger the reasoning process of DSAPE for producing recommendations according to the current level of awareness. The security awareness agent maintains this security

awareness program and reasoning process for triggering events for the higher management. Moreover, a multi-agent system can be constructed for large enterprises, where each agent monitors the awareness program of a smaller division and communicates with the rest of the agents to produce an aggregated recommendation report. **Fig. 1** illustrates the software layers of DSAPE.

The agent's developer GUI consists of six tabs. The first tab is the agent's *view*. It contains the agent's knowledge base – where the latest changes are indicated with red color; the agent's output – where messages for the reasoning process and communication with other agents are reported; agent's input – where new events can be indicated; and agent's connection – where the agent can connect to other agents and exchange information. The second tab is the agent's *theory*. It consists of rules in Jess, which describe the composition of the security awareness program's modules and metrics, the reasoning process for local recommendations and the communication with other agents. The third tab is the agent's *facts*. They are the basic definitions and facts declarations in Jess that are used in the reasoning process. The fourth tab is the agent's *model*. It traces the latest reasoning process of the extended DECKT. The fifth tab is the agent's *recommendations*. It summarizes the active recommendations of DSAPE along with the remaining budget. The recommendations are grouped in four categories (periodic, temporal, casual and reactive), as described in sub-section 3.3. The last tab is the *DSAPE tab*. It illustrates the local security awareness program as well as its modules and metrics along with their evaluation, their weights (both for individual metrics and their corresponding module weights), as well as their cost values. Low awareness values (less than 30%) are marked with red color, neutral awareness values (30%-70%) are marked with blue color and high awareness values are (more than 70%) are marked with green color. The total cost is marked with red color when it reaches a cost threshold, which is specified by the user (as percentage of the organizations budget), to denote the low limited capabilities for performing security awareness activities during the remaining economic period.

DSAPE features a GUI for end-users, implemented in HTML and JavaScript. Users are expected to update the metrics after an awareness event; choosing the metric to

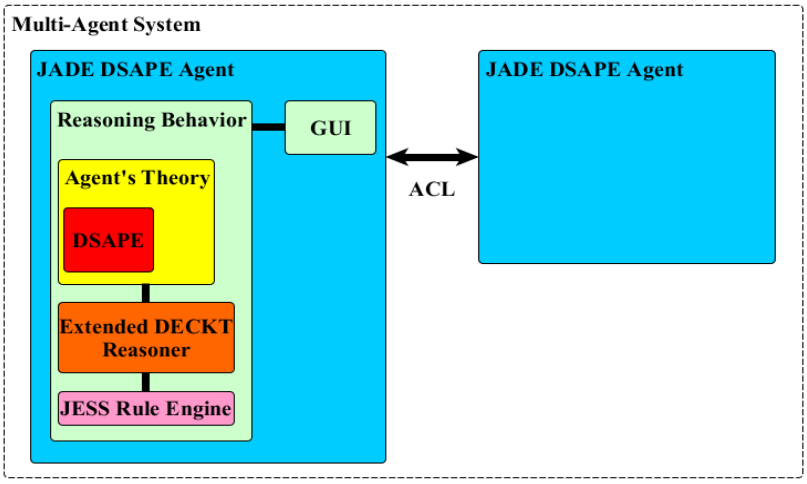

Fig. 1. Software layers of DSAPE.

`

be updated and passing the new parameters. Then the agent processes the new event and evaluates the new state of the security awareness program. Users can also view the program's details and receive the latest recommendation report.

3.2 Value aggregation

A security awareness program is composed of modules and each module is composed of evaluated metrics. When a metric is evaluated there are two functions for calculating its awareness value and cost respectively. The awareness value of a module is the weighted sum of its underlying metrics' value, while its cost is the sum of the costs of these metrics. Respectively, the awareness value of the program is the weighted summation of the program's modules and the total cost is the summation of their costs.

When a metric is evaluated, an event occurs to determine the new program state. The metric's functions for calculating its security awareness level and its cost are executed. The event can change the awareness and cost values of the relevant metric, module and program and can trigger the reasoning process of the awareness agent. Metrics and modules can be deployed or deleted dynamically along with their weights and evaluating functions. The weights can then be re-distributed automatically, or, ideally, via user interaction. The DSAPE application includes 12 modules and 18 metrics, as previously presented, but can be tailored to each organization's needs. For example, an organization can begin with only a small subset of the proposed metrics and modules and add more of them as the security awareness evaluation proceeds. Moreover, it can construct its own metrics and modules according to its business practices and the latest progresses in security awareness.

3.3 Recommendation Reports

The recommendation report procedure is one of the main features of DSAPE. We can model recommendations for periodic, temporal, casual and reactive actions. *Periodic* actions include events that have to be performed in a recurrent manner. Recommendations for periodic actions can include notations for annual and monthly events (e.g. annual surveys and monthly newsletters). A rule is implemented for every periodic action which is triggered after the last update of the relevant metric and determines when the new update has to be performed according to the action's period. *Temporal* actions are operations that must be performed at a specific time. For example, the immediate informing of the company's personnel for a serious security incident via an iNotice (and then re-evaluate the security awareness program by performing the relevant metrics). A rule is modeled to erase the notation once the action is performed. *Casual* actions are occasional procedures that are assigned by the management team. An unplanned security audit at a department could be a casual action. A rule is in charge of reporting and erasing the relevant notification. *Reactive* actions are automated activities that are planned by the security management team and determine the organization's reaction based on the current state of the security awareness program. It is a core AI process that performs the organization's strategy for improving the

security awareness level. As an example, consider an organization with low security awareness level. First, it should initiate actions to train its employees. Then, it should evaluate the training procedure. If it is acceptable, the organization should proceed to a sampling security audit to its departments. If the audit accents a low enforcement of the security procedures that had been communicated during training, the organization should plan a new training action. Moreover, the reactive plan can take into account an annual budget for security awareness activities. The cost of deploying the periodic actions and the rest actions the have been performed so far is abstracted, and reactive actions are suggested based on the remaining budget.

DSAPE can be used to estimate the effectiveness of an implemented metric through time by tracing its past values. It can also be used to identify the most efficient security awareness operations based on this effectiveness control and a cost-benefit procedure. The framework can, thus, indicate preferable metrics for a specific category and which should be avoided or even eliminated. An analysis of these reports by the management team can lead to a better adjustment of the security awareness program's modules, metrics and their relevant weights for this specific organization.

3.4 Multi-agent DSAPE application

The single agent DSAPE application can be utilized by small and medium companies for evaluating their security awareness level, as described in the previous sections. However, large enterprises with many divisions and/or offices in various premises or with global reach cannot be effectively evaluated by a single agent.

Thus, a multi-agent DSAPE system is proposed to meet such requirements, where each different division can deploy a DSAPE agent for monitoring and improving its local security awareness level. Other than the recommendations pertinent to this local division, agents can be modeled to communicate high level information to a master agent. The master agent collects all these pieces of knowledge and presents them to the higher levels of management, located in the company headquarters or elsewhere. Thereby, conclusions can be derived about regional security awareness behavior and habits as well as the security awareness status of the organization as a whole. This knowledge can also be combined with other decision making systems (e.g. Management of Information Systems), assessing the upcoming actions in improving the overall security awareness level and the investments in specific countries or geographical regions.

The master DSAPE agent can apply more complex metrics and modules as well as a social reasoning process that runs the DSAPE multi-agent community. The overall security awareness level of a DSAPE multi-agent system is estimated by this master agent and is calculated as the weighted summation of the underlying local security awareness programs and the overall cost as the summation of the costs for evaluating these programs.

`

3.5 Demonstration

This section presents an application of a single DSAPE agent evaluating the security awareness level of a small organization with 50 employees. For the sake of simplicity, periodic, temporal and casual actions are not included.

Sample weights (as described in subsection 2.2 above) are set and the reactive strategy (described in sub-section 3.3*)* is modelled. It is assumed that all metrics have been evaluated by the organization at least once. However, the security awareness level is low (28.6%). An e-learning session is performed and its effectiveness is evaluated, updating the e-learning module (metrics M13 to M15). For M_{13}, 40 of the 50 employees visit the e-learning web site, increasing the metric's value from 40% to 80%. For M_{14}, 34 of the 40 employees that visit the e-learning site (M*13*) register to the e-learning program, achieving 85%. For M_{15}, 32 of the 34 registered employees (M_{14}) complete the program successfully, succeeding 95%. Thus, the e-learning module takes the high value 89%. The security awareness level is increased to 33.7% and the program's cost is increased by 60$. DSAPE then indicates an audit action would be beneficial, thus an audit session is performed and the result is evaluated (updating the independent observation module, metrics M4-5, accordingly). A phishing e-mail was sent to each employee (totally 50 e-mails), exhorting the receiver to visit a suspicious web site. For M_4, only 5 employees didn't visit the web site (45 successful phishing attacks), thus the metric achieves a low value of 10%. Moreover, the awareness team sent an e-mail to every employee with a link to a legitimate web site that informs the visitor about new threats and security issues. For M_5, 5 of the total 50 employees eventually visit the web site to get informed about the latest news in security, accomplishing 10%. The pure performance of the personnel reflects to the low value for the independent observation module of only 10%. The security awareness level is decreased to 32.3% and the cost is increased by 30$. DSAPE reveals the low enforcement of the security practices that were learned and suggests planning a new training activity.

Fig. 2 summarizes the security awareness program's state at two of the demonstration phases detailed above. *"A"* presents the initial state (security awareness level is low and denoted with red color), while *"B"* is the final state. The corresponding metric and module weighs as well as cost values can also be seen in this figure.

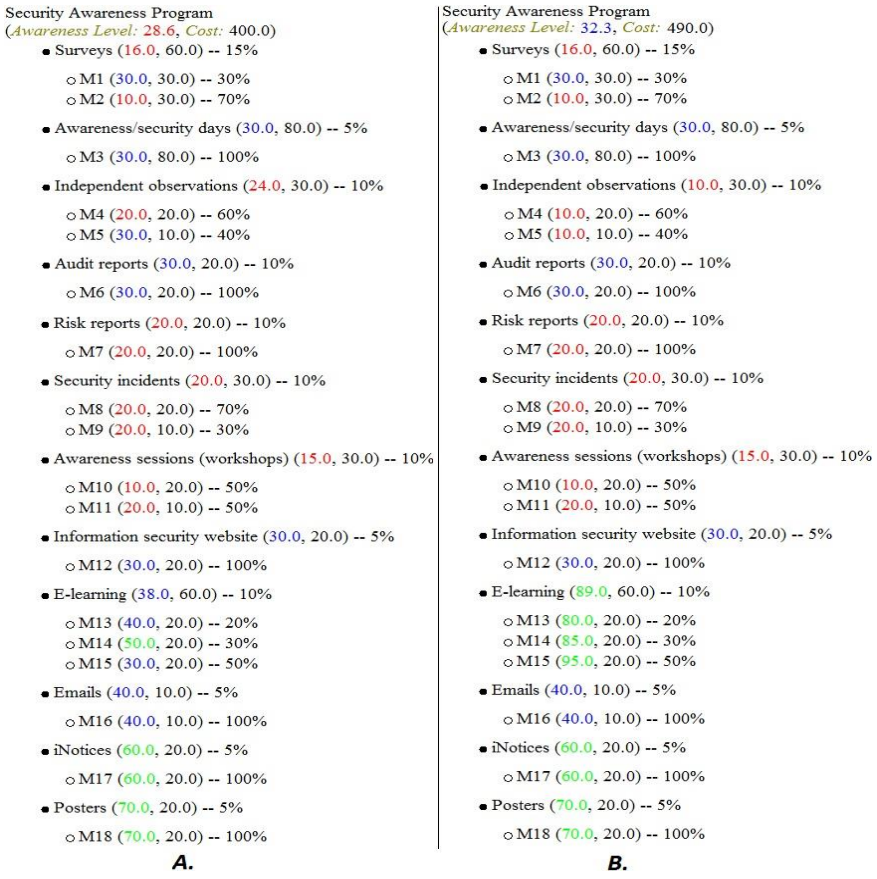

Fig. 2. The single DSAPE agent security awareness program.

4 Conclusion

The key to a successful awareness program is continuous monitoring and improvement, which can only be proven by applying and following a specific effectiveness measurement approach. Through the use of DSAPE, the evaluation methodology and accompanying model-based tool proposed in this work, the awareness team and higher management will have a dynamic tool providing awareness evaluation and monitoring. DSAPE can, thus, be utilized to provide an assessment and validation of the results of a deployed program, enabling the stakeholders to monitor the level of the program's success with regard to meeting their initial targets and its effect on the organization's actual security awareness culture. The awareness group can then make informed decisions, setting targets for the next awareness program based on the results drawn from the current program's evaluation. Such an ongoing evaluation will provide the means to take corrective actions to ensure the best possible result for their effort and investment.

Acknowledgements. This work was funded by the General Secretarial Research and Technology (G.S.R.T.), Hellas under the Artemis JU research program nSHIELD (new embedded Systems arcHItecturE for multi-Layer Dependable solutions) project. Call: ARTEMIS-2010-1, Grand Agreement No: 269317.

`

5 References

1. K. Rantos, K. Fysarakis and C. Manifavas, “How effective is your security awareness program? – An evaluation methodology”, Information Security Journal: A Global Perspective, 21:6, 328-345, 2012.
2. T Tryfonas, E Kiountouzis, A Poulymenakou, "Embedding security practices in contemporary information systems development approaches", Information Management & Computer Security, Information Management & Computer Security, Vol. 9 Iss: 4, pp.183 - 19, 2001.
3. N.E. Petroulakis, I.G. Askoxylakis, T. Tryfonas, "Life-logging in smart environments: Challenges and security threats," Communications (ICC), 2012 IEEE International Conference on , vol., no., pp.5680,5684, 10-15 June 2012.
4. Deloitte, “Global Security Survey”, 2010.
5. European Network and Information Security Agency (ENISA), “The new users’ guide – How to raise InfoSec Awareness”, 2010.
6. National Institute of Standards and Technology (NIST), "Special Publication 800-50: Building an information technology security awareness and training program", 2003.
7. H. A. Kruger and W. D. Kearney, “A prototype for assessing information security awareness,” Computers & Security, Elsevier, vol. 25, pp. 289-296, 2006.
8. Savola R., "A Novel Security Metrics Taxonomy for R&D Organizations", in proceeding of: The ISSA 2008 Innovative Minds Conference, ISSA 2008, Gauteng Region (Johannesburg), South Africa, 7-9 July 2008.
9. National Institute of Standards and Technology (NIST), "Special Publication 800-55, Revision 1: Performance Measurement Guide for Information Security", 2008.
10. National Institute of Standards and Technology (NIST), "Special Publication 800-16: Information technology security training requirements: a role- and performance-based model", 1998.
11. E. T. Muller, “Commonsense reasoning,” M. Kaufmann, 2010.
12. T. Patkos and D. Plexousakis, “DECKT: epistemic reasoning for ambient intelligence,” ERCIM News magazine – Special Theme: Intelligent and Cognitive Systems, issue 84, January 2011. http://ercim-news.ercim.eu/en84/special/deckt-epistemic-reasoning-for-ambient-intelligence .
13. Oracle-Java, JESS: the Rule Engine for the Java Platform. Available at: http://herzberg.ca.sandia.gov/ .
14. G. Hatzivasilis, “Multi-agent distributed epistemic reasoning in ambient intelligence environments,” Master Thesis, University of Crete, Computer Science Department, Greece, Crete, Heraklion – Foundation for Research and Technology – Hellas, Institute of Computer Science (FORTH-ICS), November 2011. http://www.ics.forth.gr/_publications/Hatzivasilis_Master_Thesis.pdf .
15. JADE, Java Agent DEvelopnet (JADE) Framework. Available at: http://jade.tilab.com/ .
16. FIPA-ACL, Agent Communication Language (ACL). Available at: http://en.wikipedia.org/wiki/Agent_Communication_Language .
17. H. A. Kruger and W. D. Kearney, “A prototype for assessing information security awareness,” Computers & Security, Elsevier, vol. 25, pp. 289-296, 2006.