Miguel, J. [et al.] (2014) Towards a normalized trustworthiness approach to enhance security in on-line assessment. *2014 Eighth International Conference on Complex, Intelligent, and Software Intensive Systems, CISIS 2014, 2-4 July 2014, Birmingham, United Kingdom: proceedings.* [S.l.]: IEEE, 2014. Pp. 147-154 Doi: http://dx.doi.org/10.1109/CISIS.2014.22.

Towards a Normalized Trustworthiness Approach to Enhance Security in On-line Assessment

Jorge Miguel[1], Santi Caballé[1], Fatos Xhafa[1], Josep Prieto[1], Leonard Barolli[2]
Department of Computer Science, Multimedia, and Telecommunication
Open University of Catalonia
Barcelona, Spain
{jmmoneo, scaballe, fxhafa, jprieto}@uoc.edu
[2] Fukuoka Institute of Technology, Department of Information and Communication Engineering
Fukuoka, Japan
barolli@fit.ac.jp

***Abstract*—This paper proposes an approach to enhance information security in on-line assessment based on a normalized trustworthiness model. Among collaborative e-Learning drawbacks which are not completely solved, we have investigated information security requirements in on-line assessment (e-assessment). To the best of our knowledge, security requirements cannot be reached with technology alone; therefore, new models such as trustworthiness approaches can complete technological solutions and support e-assessment requirements for e-Learning. Although trustworthiness models can be defined and included as a service in e-assessment security frameworks, there are multiple factors related to trustworthiness which cannot be managed without normalization. Among these factors we discuss trustworthiness multiple sources, different data source formats, measure techniques and other trustworthiness factors such as rules, evolution or context. Hence, in this paper, we justify why trustworthiness normalization is needed and a normalized trustworthiness model is proposed by reviewing existing normalization procedures for trustworthy values applied to e-assessments. Eventually, we examine the potential of our normalized trustworthiness model in a real online collaborative learning course.**

Keywords- trustworthiness; normalizarion; e-assessment; information security; collaborative learning

I. INTRODUCTION

Recent works in e-Learning security have shown that information security in e-Learning must be managed with complementary trustworthiness approaches [1]. Although online assessments (e-assessments) in both continuous evaluation and collaborative learning have been widely adopted, there exist still drawbacks which limit their potential, among them, information security in e-assessment [2], [3]. In these previous works we discussed how security requirements in e-assessment cannot be reached with technology alone and a trustworthiness model for the design of secure learning assessment is proposed.

In the scope of our research, we consider e-assessment following the evaluation model used in UOC courses. Evaluation models used in UOC may be classified in accordance with several factors such as type of subjects, specific evaluation model, evaluation application or agents involved in the evaluation processes. Namely, in manual evaluation methods, tutors usually participate directly and intensely in the evaluation process but this model has scalability problems. On the other hand, although automatic methods do not involve tutors participation, the model does not carry out desirable integrity levels. Therefore, our assessment model proposes hybrid methods are a trade-off combination which can provide a balance between the degree of interaction and security requirements.

Regarding trustworthiness in the context of e-Learning, according to [4], a trustworthy e-Learning system is a trust-based learning system which contains reliable learning resources. Although several trustworthiness models have been defined and included as a complementary service in e-assessment security frameworks [5], modeling trustworthiness involves multiple complex and heterogeneous factors, such as data coming from different sources and with different formatting, which cannot be managed without normalization. Some authors have proposed normalized trustworthiness approaches focused on specific trustworthiness features [6], [1] in order to combine multiple trustworthiness sources and to manage different resources. However, to the best of our knowledge, a normalization method dealing with e-assessment, should include a more complex set of factors related to specific security properties in e-assessment presented in [3] by examining and selecting most relevant security properties in order to build a secure e-assessment model.

Therefore, in this paper, we discuss trustworthiness multiple data sources, different data formats, trustworthiness measurement techniques and other factors, such as trust rules, evolution or context. And then, with the aim to enhance trustworthiness analysis, we consider these trustworthiness factors and their relations in order to create a methodological approach to build our trustworthiness-based normalized model devoted to enhance security in on-line assessment.

The paper is organized as follows. Section II presents the background and context information on trustworthiness general models, e-Learning approaches and normalization proposals. A methodological approach to build our trustworthiness-based normalized model is presented in section III. A real online course is proposed in section IV to evaluate a hybrid evaluation system supported by our normalized trustworthiness model through a statistical analysis. Finally, Section V concludes the paper highlighting the main findings and outlining ongoing and future work.

II. BACKGROUND

In this section we review main works in the literature on general trustworthiness models, e-Learning approaches and normalization proposals.

A. *Trustworthiness and Security for e-Learning*

According to [7], problems encountered in ensuring modern computing systems cannot be solved with technology alone. Early research works about trustworthiness management models [6], [1] suggest that soft security, such as social control or users reputation in distributed systems, have to be used to provide information security improvements. In [6], the authors propose a social control system devoted to manage security issues when participants themselves are responsible for the security, as opposed to security implemented by external technological solutions. This model is developed by using reputation agents, which manage what information is transmitted to the other actors. The authors in [1] explained why traditional network security mechanisms are incomplete in their function to manage trustworthiness, and a general model based on recommendations is provided. The authors point out that information security solutions need more effective trust complementary management schemes and techniques; in this work the following three trends in current security practice impact are presented: (i) Hard Security based on Public Key Infrastructures (PKI), which is a relevant example of hard security mechanisms though hard security mechanisms do not say anything about trustworthiness; (ii) Centralized Protocols based on protocols, which use a common Trusted Authority (TA), to form a trust relationship between two mutually distrusting entities, and considers that a TA can never be a good enough recommender for everyone in a large distributed system making its credibility depleting, and its recommendations increasing in uncertainty, whilst its community of trustees grows; (iii) Implicit Trust Assumptions, which assumes that if a secure system is desired, trust assumptions must be explicit and qualification is required, for instance, under what circumstances trustworthiness has been defined [1].

More recently, in [8],[9], it is discussed that security is both a feeling and a reality. The author points out that the reality of security is mathematical based on the probability of different risks and the effectiveness of different countermeasures. On the other hand, the authors state that security is also a feeling, based not only on probabilities and mathematical calculations but on psychological reactions to both risks and countermeasures [9]. Since this model considers two dimensions in security (i.e. feelings and reality) and being aware that absolute security does not exist, it can be concluded that any gain in security always involves a trade-off in this context, where technological security and trustworthiness analysis are required as a complete and hybrid model devoted to ensure e-Learning [2].

B. *Trustworthiness General Models*

In information technologies environment, the first formally trustworthiness model was proposed in [10] from three levels, as follows: (i) Basic trust is the general trusting disposition of an agent A at time T; (ii) General trust represents the trust that agent A has on agent B at time T; (iii) Situational trust is the amount of trust that an agent A has in another taking into account a specific situation.

With the purpose to measure trustworthiness levels, in [11], the author designs a survey to explore interpersonal trust in work groups, identifying trust-building behaviors ranked in order of importance. These behaviors can be used as trustworthiness factors, which can measure trust and they are classified as Trustworthiness Building Factors (TBF) and Trustworthiness Reducing Factors (TRF). Moreover, according to [4], there are different aspects of considering on trust, different expressions and classifications of trust issues. In essence, we can summarize the aspects presented in this work defining the following rules: (i) Asymmetry, A trust B is not equal to B trust A; (ii) Time factor, trustworthiness is dynamic and may evolve over the time; (iii) Limited transitivity, if A trusts C who trusts B then A will also trust B, but with the transition goes on, trust will not absolutely reliable; (iv) Context sensitive, when context changes, trust relationship might change too.

The authors in [12] proposed a data provenance trust model, which takes into account various factors that may affect trustworthiness and, based on these factors, this model assigns trust scores to both data and data providers. These scores represent key information and users may decide whether to use the data and for what purposes. The trust score of an item is computed by taking into account four factors: (i) data similarity, the likeness of different scores in the same set; (ii) path similarity, regarding intermediate agents that processed data from source to destination; (iii) data conflict, inconsistent descriptions or information about the same entity; and (iv) data deduction, if the source information or the responsible party is highly trusted, the resulting data will also be highly trusted.

Software components related to trustworthiness modules have been developed recently, among them FeelTrust [13] is an application for smartphones that automatically monitors a user’s overall trustworthiness levels. To this end, FeelTrust classifies users as trusted or not depending on their interests and pair this result with feedbacks from an embedded reputation system. As stated by the authors, the FeelTrust implementation demonstrates the feasibility of security tasks using trustworthiness models. This approach is based on two modules oriented to collect trustworthiness data: Monitor

Behavior, module that monitors and collects sensor data and Manage Feedbacks, module that manages feedbacks (i.e. scores, ratings or recommendations).

Resource Description Framework[1] (RDF) is a standard model for data interchange on the Web, in [14] it is proposed a trust model for RDF data that considers trustworthiness on the level of data sources. This model is devoted to enable a trust infrastructure for the Web by developing concepts for automatic trust assessment based on provenance information and on the opinion of other information users. Furthermore, this approach provides trust-aware data access methods and concepts to implement trust-aware systems.

C. *Trustworthiness in e-Leaning*

Regarding trustworthiness and e-Learning, according to [4], a trustworthy e-Learning system is a trust-based learning system which contains reliable serving peers and useful learning resources.

In [15] it is presented a service platform for mobile learning with trustworthy service provisioning based on an integration of grid services, on demand e-Learning, and trusted mobile asset tracking. The service platform, called MiQ-SP [15], is designed for mobile learning with trustworthy service provisioning and developed based on an integrated service network concept, for tracking e-Learning participants and managing e-Learning assets.

The studies presented in [5], [16] stem from the difficulties of guaranteeing the quality and trustworthiness of learning resources and participants. These drawbacks make learners and educators have no enough confidence in participating in web-based learning [5] and it may be difficult at times for users to select the most appropriate content themselves, in order to enhance their learning experience [16]. In [5] a quality assurance and trustable e-Learning environment with quality certification and trust evaluation is proposed. This model is based on a Service-oriented architecture and combines static quality certification and dynamic trust evaluation. The trustworthiness approach presented in [16] is based on trustworthiness, similarity and knowledge gains. As stated by the authors this model provides an effective solution to support peer-based information sharing within web-based contexts.

In [17] a peer-to-peer based social network to enhance the quality of e-Learning is presented. This network is based on knowledge sharing in virtual learning communities. In order to organize and provide better resource management, each peer has to classify content and evaluate its quality (i.e. rating of the resource), number of times the site is accessed, and the matching degree [17].

D. *Normalized Trustworthiness Models*

The concept of normalized trustworthiness is introduced in [18] as a trust relationship in terms of a vector which is normalized by a trust policy. Each element, in trust vector, represents a parameter that contributes towards the trust evaluation in a specific time and context (i.e. A trusts B at a time T and for a particular context C). The normalization model is based on the trust policy vector which is a vector that has the same dimension as the trust vector; the elements are real numbers in the range [0, 1] and the sum of all elements is equal to 1. Hence, each normalized value is the result of a trustworthiness value multiplied by the corresponding weight in the policy vector.

Further works [19],[20],[21] improve the normalization process by proposing more complex functions; as stated in [19], a simple arithmetic average would perform a rough compensation between high and low trustworthiness values. In order to fill this drawback, the authors propose the Weighted Ordered Weighted Averaging (WOWA), which uses two sets of weights: one corresponds to the relevance of the sources, and the other corresponds to the relevance of the trustworthiness values. According to [21] a light-weight mathematical model can be used to represent the collected evidences. Following this model, all the trustworthy evidences are represented as a vector and then the trustworthy values can be represented as a trustworthy matrix where each row of matrix is a trustworthiness item, and each column is a trustworthiness value. Moreover, the preferences of trustworthy evidences are represented as a weight vector and, after normalization; the trustworthiness of each user can be evaluated as the normalized matrix (i.e. multiplication between a trustworthiness matrix and a weight vector).

Another perspective is focused on normalize trustworthiness values by subtracting the average value for a user, in [22], the concept of Filler Mean Target Difference (FMTD) attribute is introduced as the overall average subtracted from a specific rating as a normalizing factor. This technique is also proposed in [20] devoted to normalize the vectors representing the profiles of items by the utility value. Finally, in [23] it is presented how to estimate information trustworthiness by considering multiple information sources. This problem is formulated as a joint matrix factorization procedure where different sets of users from different sources are partitioned into common groups and rating behavior of groups is assumed to be consistent across sources.

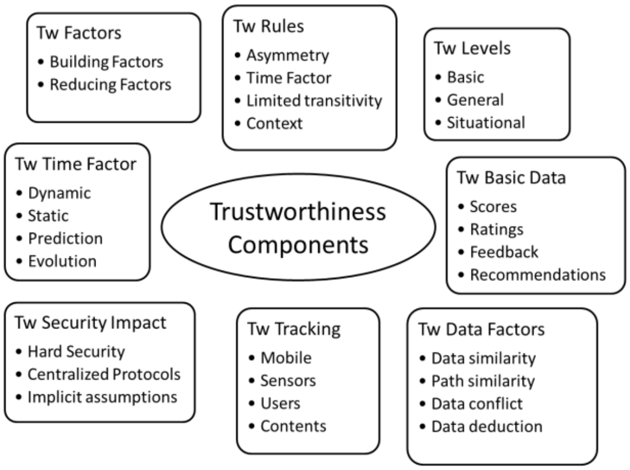

Figure 1. Trustworthiness components

[1] RDF specification is found at: http://www.w3.org/RDF/

To sum up, trustworthiness components, features and factors which have been identified in these sections, from previous works about trustworthiness models and normalization are summarized in Fig. 1.

III. Normalized Trustworthiness Model

We present a methodological approach to build our normalized trustworthiness-based model. This proposal is based on our previous research work presented in [2],[3], by enhancing normalization properties which are presented in this section.

To this end, we first analyze trustworthiness data sources, then, normalization functions are presented and, finally, the normalized trustworthiness levels are proposed.

A. *E-Assessment Data Sources*

Our model is based on an input set formed by multiple trustworthiness data sources. Due to each Trustworthiness Data Source (TDS) has different formats and characteristics, the foremost step regarding normalization, is to identify Trustworthiness Data Format (TDF) generated by each TDS. A TDS is implemented via technological tools integrated in the learning process intended to collect quantitative trustworthiness data. These tools are summarized as follows:

- Ratings. Students can share many objects (e.g. documents, notes, folders, blogs, etc.), most of them can be rated in terms of qualification. Therefore, we select the most usual objects that can be rated and offer a high level of trustworthiness data; namely, notes (posts on forums) and shareable documents. The document's quality can be rated by selecting one of the ratings offered by the LMS and the TDF follows a typical 5-point Likert scale.
- LMS usage indicators. Besides ratings, we collect students' general activity in the LMS. Overall values, such as number of dynamic items (i.e. forum posts) created by a student, are automatically collected.
- Student's reports. Students' reports are another source of information for trustworthiness. The reports contain questions with quantitative responses in a 1-10 scale. The coordinator of a group of students has to complete two different reports for each phase of the course. The first type is public and the group is evaluated, the second one is private (i.e. only accessible by the tutor) and contains evaluation data for every members of the group.
- Questionnaires. We propose an abstract proposal of a questionnaire, which is instanced when the model is implemented. Our questionnaire is arranged in five sections (see an excerpt of the questionnaire with the first three sections in Table 1), each section contains groups of questions of the same type (Questions and Sections, Q/S and questions evaluating the questionnaire QQ). For each question, the assessment scale is defined. The questions about the group may refer to each member of a group (Group or Individual, G/I and Scale, S). Trustworthiness building and reducing factors are also included as target measurable factors (Trustworthiness Factor, TF) [11]. Finally, the type of the question is defined, the value text (T) means open text responses where the student introduces his or her comments, and if the question is quantitative it is represented by N (number in the scale) or R (if the question is a ranking).

TABLE I. Abstract Questionnaire example Description

QS	***GI***	***TNR***	***S***	***Q/S Description***	***TF***
S1	Trustworthiness building factors				
Q1	I	N	1-5	Honest communication	TBF1
Q2	I	N	1-5	Commitments accomplishment	TBF3
Q3	I	N	1-5	Confidence in abilities	TBF2
Q4	I	N	1-5	Regard for partners' statements	TBF4
Q5	I	N	1-5	Mutual help	TBF5
Q1-5	I	T	-	Comments regarding Q1...Q5	
S2	Trustworthiness, security and reliability				
Q6	I	R	1-5	Individual trustworthiness level	TBF
Q7	I	R	1-5	Individual security level	TBF
Q8	I	R	1-5	Individual reliability level	TBF
S3	Trustworthiness reducing factors				
Q9	G	T	1-5	Concerned about individual goals	TRF1
Q10	G	T	1-5	To avoid taking responsibility	TRF3
Q11	G	T	1-5	To avoid analyzing the facts	TRF4
Q12	G	T	1-5	To make excuses	TRF5
Q13	G	T	1-5	To blame others	TRF5

B. *Normalizing Data Sources and Trustworthiness Levels*

As presented in above subsection, each TDS follows its own TDF, therefore a preliminary normalization process is needed in order to normalize these sources following an unified format. To this end, we introduce now the concept of normalized trustworthiness indicator tw_i^N (with $i \in I$, where I is the set of trustworthiness indicators) as a measure of trustworthiness factors. Trustworthiness factors have been presented as those behaviors that reduce or build trustworthiness in a collaborative group and they have been considered in the design of both questionnaires and reports. A tw_i is associated with one of the measures defined in each e-assessment instrument (i.e. ratings, questionnaires, reports, etc.) and can be represented following this expression:

$$tw_{a_{r,s}}^{N} = N\left(tw_{a_{r,s}}\right)\ a \in \{Q, RP, LI, RL\}, r \in R, s \in S$$

where Q is the set of responses in Questionnaires, RP is the analogous set in Reports, LI is the set of LMS general indicators and RL is the set of Ratings in the LMS. S is the set of students and R is the set of rules and trustworthiness characteristics (e.g. time factor).

The Normalization function $\mathrm{N}_1\left(tw_{a_{r,s}}\right)$ normalizes the trustworthiness indicator by transforming the indicator value into a unified TDF which values are between 1 and 5. Once a value has been normalized, value 1 always means a very low trustworthiness case although the indicator represents, for instance, a TRF or a risk factor. The normalization function considers the normalization new cases. In other words, as trustworthiness indicators are related to reducing and building factors, function N_1 normalizes all values as trust-building values as follows:

$$N_1(tw_i) = \begin{cases} \max\left(tw_i\right) - tw_i\,, & i \in I_R \\ tw_i, & i \in I_B \end{cases}$$

where I_R is the set of trustworthiness indicators which represent reducing trustworthiness behaviors and I_B is the set of indicators based on building factors.

In a questionnaire, a student can evaluate, in the same question, every member in the group; in order to tackle this case, we propose the following normalization function:

$$N_2\left(tw_{i,a}\right) = \sum\nolimits_{s=1}^{m} \frac{tw_{i,a}}{m-1}, \qquad \mathrm{s} \neq \mathrm{a}, i \in I_I$$

where $a$is the target student (i.e. the student which is evaluated); I_S is the set of indicators measuring individual assessments in a student group; and m is the number of students in the group of the student a.

Moreover, we need a linear transformation to convert one Likert scale to the normalized scale 1-5. We propose the Linear Stretch Method [24] as follows:

$$N_3(tw_i, p_1, p_n) = s_1 + \frac{(s_n - s_1)(tw_i - p_1)}{(p_n - p_1)}$$

where the primary scale is consecutively numbered from p_1 to p_n and $(s_1\,,\, s_n\,)$ is the target scale; p_1 and p_n are introduced as parameters because we have to manage multiple scales.

Regarding ratings, it is worth mentioning that each group has its own domain and a reference value has to be taken. To this end we normalize the number of rates that a student has done as follows:

$$N_4(tw_i) = \frac{tw_i * (p_n - p_1 + 1)}{T_G}$$

where T_G is the maximum number of ratings by an student in the group G and $(p_n - p_1 + 1)$ is the number of items in the rating scale.

The latter case which requires normalization is related to both student reports and questionnaires, as aforementioned public reports evaluate the group and each student is evaluated in public reports, although we have managed a similar case with $N_2(tw_i)$ in the case of public reports we do not have individual values. We propose to tackle this situation by estimating students' values as the group evaluation:

$$N_5(tw_s) = tw_G, \qquad \mathrm{s} \in \mathrm{G}$$

where G is the group of the student s and tw_s is an indicator from a public student report.

Finally, we apply each normalization function $N_j(tw_i)$ in order to obtain the normalized indicator $N(tw_i)$for those indicators which need normalization by the conditions (i.e. in a selective way) presented in this section.

For instance, if we calculate the normalized value of a public student report indicator, the following normalization functions are needed:

$$tw_s^N = N_4\left(N_5(tw_s)\right)$$

C. Modeling Normalized Trustworthiness Levels

The concept of trustworthiness level Ltw_i is a composition of indicators over trustworthiness rules and characteristics. For instance, we can consider two trustworthiness indicators (tw_a and tw_b). These indicators are different, the first indicator could be a rating in a forum post and the second one a question in a questionnaire; but they measure the same trustworthiness building factor (e.g. communicates honestly). Trustworthiness levels Ltw_i must be normalized; to this end, we have selected as normalization model a weight-based normalization. Following this approach, we previously need to define the weights vectors:

$$w = (w_1, \dots, w_i, \dots w_n), \sum_{i}^{n} w_i = 1$$

where n is the total number of trustworthiness indicators and w_i is the weight assigned to tw_i.

Then, we define trustworthiness normalized levels as:

$$Ltw_i^N = \sum_{i=1}^{n} \frac{(tw_i * w_i)}{n}, i \in I$$

Therefore, trustworthiness levels allow us modeling students' trustworthiness as a combination of normalized indicators using each TDS. It is important to note that a level can be composed by trustworthiness levels, that is:

$$L_i^N = \sum_{i=1}^{n} \frac{(Ltw_i^N * w_i)}{n}, i \in I$$

Furthermore, we consider the time factor as a normalization component, which also allows us to analyze both relation and similarity [3],[2]:

$$r_{a,t,tt} = \frac{\sum_{i=1}^{n}(tw_{a_{t,i}} - \overline{tw}_{a_t}) * (tw_{a_{tt,i}} - \overline{tw}_{a_{tt}})}{\sqrt{\sum_{i=1}^{n}(tw_{a_{t,i}} - \overline{tw}_{a_t})^2} * \sqrt{\sum_{i=1}^{n}(tw_{a_{tt,i}} - \overline{tw}_{a_{tt}})^2}}$$

where t is the target point in time and tt is the reference point in time (i.e. t is compared against tt), tw_a is the target trustworthiness indicator, tw_b is the second trustworthiness indicator in which tw_a is compared (i.e. similarity, correlation, anomalous behavior, etc.), $\overline{tw}_a$ and $\overline{tw}_b$ are the average of the trustworthiness indicators and n is the number of student's provided data for tw_a and tw_b indicators.

IV. Evaluation and Analysis of the Results

This section presents a real online course intended to evaluate a hybrid evaluation system supported by our normalized trustworthiness model. Firstly, the context of the course is defined. Then, we propose a subset of data collected by data sources described in the normalized trustworthiness model presented. And finally, it is presented a statistical and evaluation analysis based on our model.

A. *Real online course features*

The experiment proposed is focused on a real online course at the Open University of Catalonia, which has these main features:

- Collaborative activities represent a relevant component of the e-assessment of the course.
- Students' evaluation is based on a continuous evaluation model by using several manual evaluation instruments.
- Number of students participating: 12 students distributed in 4 groups.
- The course follows four stages (t_1, t_2, t_3, t_4) that can be taken as time references in order to evaluate and to analysis results.
- At the end of each stage, each student completes a questionnaire (Q_1, Q_2, Q_3, Q_4). These questionnaires refer to the set Q defined in the trustworthiness model section.
- Each stage is performed in collaborative work groups and is coordinated by different members one of the group.
- The coordinator of the group completes two reports (i.e. the set RP in the model, public $\overline{RP}$ and private RP) at the end of each stage evaluating the members of his or her group:

$$\left(RP_1, RP_2, RP_3, RP_4, \overline{RP}_1, \overline{RP}_2, \overline{RP}_3, \overline{RP}_4\right)$$

- General e-Learning activities are supported by a standard LMS which offers both rating systems (i.e. the set RT in the model) and general learning management indicators (LI).

B. *Defining Normalized Indicators*

In this subsection we define and calculate several trustworthiness indicators normalized by the model presented in this paper; the set of indicators have been selected with the purpose to cover all normalization functions. As first stage we take the following set of basic indicators:

$$\left(\, tw_{q1_{s1,t2}}, \dots, \, tw_{q5_{s1,t2}}\right) \dots \left(\, tw_{q1_{sN,t2}}, \dots, \, tw_{q5_{sN,t2}}\right)$$

where the data source is a questionnaire, namely, responses over (1..5), (i.e. one indicator per response); $t2$ represets the time factor fixed on stage 2 of the online course; and $s1..N$ are the students in the course.

Following the model, we have to determine the normalization functions required. In this case, $tw_{qi_{t2}}$ represents a student response referred to a group of students. Therefore, in order to tackle this case, $N_1(tw_i)$ function must be applied. Further normalization functions are not needed in this case, hence the normalized indicator results as follows:

$$tw_a^N = N_1\left(\, tw_{q1\dots5_{s1,t2}}^N, \dots, \, tw_{q1\dots5_{sN,t2}}^N\right)$$

For the sake of simplicity, we rename this indicator as tw_a^N omitting students, questions and time parameters (i.e. $q_{i_{s1,t2}}$). Note that tw_a^N is a list of indicators.

Similarly, we define the following additional indicators:

$$tw_b^N = N_3\left(N_5\left(\, tw_{RT_{j,s,t2}}\right)\right), j \in (1..19), s \in (1..12)$$

$$tw_c^N = N_3\left(\, tw_{\overline{RT}_{j,s,t2}}\right), j \in (1..19), s \in (1..12)$$

$$tw_d^N = N_3\left(N_1\left(\, tw_{q_{j,s,t2}}\right)\right) \; j \in (9..16), s \in (1..12)$$

where tw_b^N represents student's scores for each question j in the private report $RT_{t=2}$; tw_c^N represents student's scores for each question j in the public report $\overline{RT}_{t=2}$; and tw_d^N is the list of indicators that represent each group's evaluation for questions 9…16 in the questionnaire.

Finally, we define two indicators related to ratings and LMS general data:

$$tw_e^N = \; tw_{RL_{j,s,t2}} \; j \in (1..n_s), s \in (1..12)$$

$$tw_f^N = N_4\left(\, tw_{LI_{j,s,t2}}\right) \; j \in (9..16), \; s \in (1..12)$$

where $tw_{RL_{j,s,t2}}$ represents the list of student's ratings in $LR_{t=2}$, n_s is the number of ratings which s has received, and $tw_{LI_{j,s,t2}}$ is the list of items created in the LMS by the student s until $t2$.

C. Statistical Analysis and Evaluation

Following the indicators defined in above section, we define the following basic trustworthiness levels:

$$Ltw_a^N = \sum_{i=1}^{n} \frac{(tw_{a,i}^N * w_i)}{n}, \qquad w_i = 1/n, \forall\, i \in I$$

where $tw_{a,i}^N$ represents each indicator in the list of indicators tw_a^N. All weights are the same value.

Fig. 2 shows each student's result. Although the scale is 1-5, we have represented as 0 those cases in which a student (i.e. s_{11}) has not been evaluated because he or she has abandoned the learning activity.

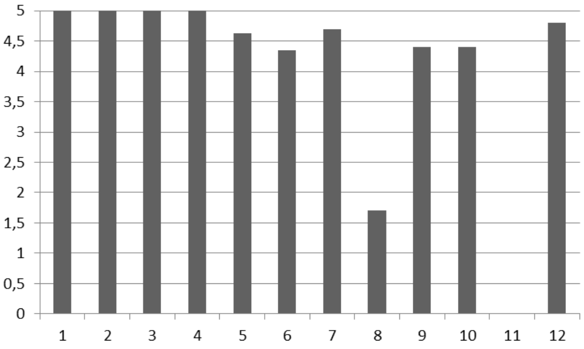

Figure 2. Trustworthiness Level: Ltw_a^N

Following the same procedure, we calculate the trustworthiness levels for a subset of the indicators:

$$Ltw_a^N, Ltw_b^N, Ltw_c^N, Ltw_d^N$$

Once the complete set of levels is available, we can compose a new complex level from them and normalizing by different weighs:

$$L^N = \sum_{i=1}^{n} \frac{(Ltw_i \;\; * w_i)}{n}$$

where $i \in (\mathrm{a,b,c,d})$ and $w = (0.2, 0.2, 0.3, 0.3)$.
Fig. 2 shows each complex level for each student.

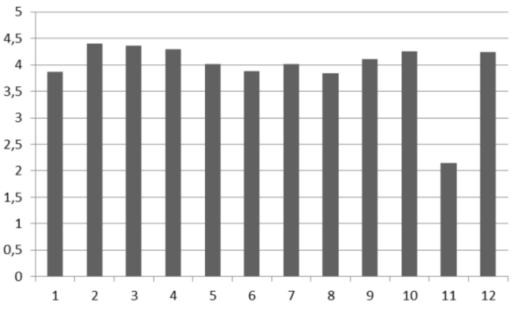

Figure 3. Complex Trusworthiness Level: L^N

In order to evaluate the model, we present two examples based on the coefficient $r_{a,\mathrm{x,y}}$. The first case is composed by calculating the similarity between the trustworthiness levels calculated for public and private students' reports, that is, tw_b^N and tw_c^N:

$$r_{\,tw_b^N,\, tw_c^N} = 0{,}90455$$

The value of r confirms that there exists linear relation between the results in public and private reports.

The second case is composed by calculating the similarity between the trustworthiness level L^N in t_1 and t_2:

$$r_{\,L^N, t_1, t_2} = 0{,}93931$$

This coefficient allows us both to predict future values and to detect anomalous students' evaluations. In this case, trustworthiness levels also follow. Fig. 4 shows the dispersion function for $r_{\,L^N, t_1, t_2}$.

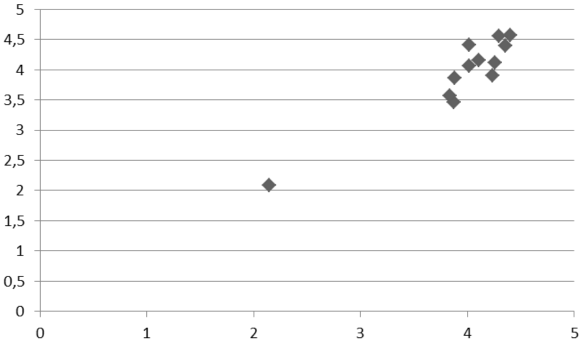

Figure 4. Dispersion function for $r_{\,L^N, t_1, t_2}$

Finally, in order to validate the hybrid evaluation model proposed, we compare our trustworthiness level with the manual and overall evaluation of the course for each student:

$$r_{\,L^N, \mathrm{ManualEval}} = 0{,}84852$$

As indicated by r, overall evaluation process and the trustworthiness level presented follow a linear relation.

V. Conclusions and Further Work

In this paper, we first have motivated the need to improve information security in e-Learning and in particular in e-assessment. Then, we have shown the feasibility of building security hybrid models, based on trustworthiness approaches. However, trustworthiness analysis in e-Learning requires normalization processes in order to tackle several trustworthiness modeling problems presented in the paper. To this end, and as a main contribution of this paper, we have proposed a methodological approach to build a normalized trustworthiness model. Finally, we have used a real online course intended to evaluate a hybrid evaluation system supported by our normalized trustworthiness model.

The experimental results showed the feasibility of modeling security by analyzing normalized trustworthiness levels and indicators. Namely, from the results comparing manual evaluation and trustworthiness levels, it can be inferred that it is viable to enhance security in e-assessment by modeling and normalizing trustworthiness behaviors.

In our future work, we would like to improve our approach in order to predict both trustworthiness students' behavior and evaluation alerts such as anomalous results. To this end, we plan to evaluate neural networks and data mining models by designing a methodological approach to construct a trustworthiness normalized model. We believe this model will represent a reference for other researchers in the domain to understand the information needed for trustworthiness purposes, and collect, normalize and analyze the data properly.

REFERENCES

[1] A. Abdul-Rahman and S. Hailes, "Using Recommendations for Managing Trust in Distributed Systems," in *Proceedings of the IEEE Intl. Conference on Communication, Malaysia*, 1997.

[2] J. Miguel, S. Caballé, F. Xhafa, and J. Prieto, "A Massive Data Processing Approach for Effective Trustworthiness in Online Learning Groups," *Concurrency and Computation: Practice and Experience*, 2014. Submitted.

[3] J. Miguel, S. Caballé, F. Xhafa, and J. Prieto, "Security in Online Assessments: Towards an Effective Trustworthiness Approach to Support e-Learning Teams," presented at the The 28th IEEE International Conference on Advanced Information Networking and Applications (AINA-2014), Victoria, Canada, 2014.

[4] Y. Liu and Y. Wu, "A Survey on Trust and Trustworthy E-learning System," 2010, pp. 118–122.

[5] Y. Liu, D. Chen, and J. Sun, "A trustworthy e-learning based on trust and quality evaluation," in *E -Business and E -Government (ICEE), 2011 International Conference on*, 2011, pp. 1–4.

[6] L. Rasmusson and S. Jansson, "Simulated social control for secure Internet commerce," in *Proceedings of the 1996 workshop on New security paradigms*, New York, NY, USA, 1996, pp. 18–25.

[7] M. J. Dark, *Information assurance and security ethics in complex systems: interdisciplinary perspectives*. Hershey, PA: Information Science Reference, 2011.

[8] B. Schneier, *Beyond fear : thinking sensibly about security in an uncertain world*. New York: Copernicus Books, 2003.

[9] B. Schneier, "The psychology of security," in *Proceedings of the Cryptology in Africa 1st international conference on Progress in cryptology*, Berlin, Heidelberg, 2008, pp. 50–79.

[10] S. P. Marsh, "Formalising Trust as a Computational Concept," University of Stirling, 1994.

[11] P. Bernthal, "A survey of trust in the workplace," HR Benchmark Group, Pittsburg, PA, Executive Summary, 1997.

[12] C. Dai, D. Lin, E. Bertino, and M. Kantarcioglu, "An Approach to Evaluate Data Trustworthiness Based on Data Provenance," in *Secure Data Management*, vol. 5159, W. Jonker and M. Petković, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 82–98.

[13] G. Carullo, A. Castiglione, G. Cattaneo, A. D. Santis, U. Fiore, and F. Palmieri, "FeelTrust: Providing Trustworthy Communications in Ubiquitous Mobile Environment," *2013 IEEE 27th International Conference on Advanced Information Networking and Applications (AINA)*, vol. 0, pp. 1113–1120, 2013.

[14] O. Hartig, "Trustworthiness of Data on the Web," presented at the STI Berlin CSW PhD Workshop, Humboldt-Universität zu Berlin, German, 2008.

[15] Z. Luo and T. Zhang, "A Mobile Service Platform for Trustworthy E-Learning Service Provisioning," in *Architectures for Distributed and Complex M-Learning Systems*, S. Caballé, F. Xhafa, T. Daradoumis, A. A. Juan, Z. Luo, and T. Zhang, Eds. IGI Global, 2009.

[16] J. Champaign and R. Cohen, "Modeling Trustworthiness of Peer Advice in a Framework for Presenting Web Objects that Supports Peer Commentary," in *Proceedings of the 20th Conference on User Modeling, Adaptation, and Personalization*, Montreal, Canada, 2012.

[17] S. J. H. Yang, I. Y. L. Chen, Kinshuk, and N.-S. Chen, "Enhancing the Quality of E-Learning in Virtual Learning Communities by Finding Quality Learning Content and Trustworthy Collaborators," *Educational Technology & Society*, vol. 10, no. 2, pp. 84–95, 2007.

[18] I. Ray and S. Chakraborty, "A Vector Model of Trust for Developing Trustworthy Systems," in *Computer Security – ESORICS 2004*, vol. 3193, P. Samarati, P. Ryan, D. Gollmann, and R. Molva, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 260–275.

[19] E. Damiani, S. De Capitani di Vimercati, P. Samarati, and M. Viviani, "A WOWA-based Aggregation Technique on Trust Values Connected to Metadata," *Electron. Notes Theor. Comput. Sci.*, vol. 157, no. 3, pp. 131–142, 2006.

[20] A. Rajaraman and J. D. Ullman, "Recommendation Systems," in *Mining of Massive Datasets*, Cambridge: Cambridge University Press, 2011.

[21] M. Li, Z. Hua, J. Zhao, Y. Zou, and B. Xie, "ARIMA Model-Based Web Services Trustworthiness Evaluation and Prediction," in *Service-Oriented Computing*, vol. 7636, C. Liu, H. Ludwig, F. Toumani, and Q. Yu, Eds. Springer Berlin Heidelberg, 2012, pp. 648–655.

[22] B. Mobasher, R. Burke, R. Bhaumik, and C. Williams, "Toward Trustworthy Recommender Systems: An Analysis of Attack Models and Algorithm Robustness," *ACM Trans. Internet Technol.*, vol. 7, no. 4, 2007.

[23] L. Ge, J. Gao, X. Yu, W. Fan, and A. Zhang, "Estimating Local Information Trustworthiness via Multi-source Joint Matrix Factorization," in *Data Mining (ICDM), 2012 IEEE 12th International Conference on*, 2012, pp. 876–881.

[24] T. Jonge, R. Veenhoven, and L. Arends, "Homogenizing Responses to Different Survey Questions on the Same Topic: Proposal of a Scale Homogenization Method Using a Reference Distribution," *Social Indicators Research*, May 2013.