Usable Trust: Grasping Trust Dynamics for Online Security as a Service

Shenja van der Graaf[1(✉)], Wim Vanobberghen[1], Michalis Kanakakis[2], and Costas Kalogiros[2]

[1] iMinds-SMIT Vrije Universiteit Brussel, Pleinlaan 9, 1050 Brussels, Belgium
shenja.vandergraaf@iminds.be, wim.vanobberghen@vub.ac.be
[2] AUEB, Athens, Greece
{kanakakis,ckalog}@aueb.gr

Abstract. This paper aims to unravel the intricacies of the mechanisms of trust vis-à-vis ICTs and the contextual logic guiding user deployment and experience, necessitating a view of trust in the digital realm as a dynamic process. Trust models tend to highlight 'well-placed trust' in their focus on drawing out (sub)components of (perceived) trustworthiness as attributes of the trusted system or party from the trustor's stance. However, less attention has been given to the trustworthy attributes, or behavior of the trusted actor. Therefore, this paper sets out to explore this linkage between ICTs and different trust-related user experiences guided by different sets of trustor attributes. In order to explore the conceptual dynamics, a two-step approach is deployed. On the basis of empirical data attention is drawn to trust levels and user segments. Preliminary insights are yielded into the trustors' segmentation validity and trust estimation accuracy by performing a small-scale experiment in the context of a fictitious online security service.

Keywords: Design · ICT · Trust drivers · Segmentation · Trustor attributes · Security

1 Introduction

Understanding why people trust and how that trust shapes social relations, has been a central interest in various scholarly domains. Social conceptualizations of trust tend to be associated with terms such as honesty, integrity and reliability; or, the extent people have 'faith in others'. Long-standing academic traditions have aimed to provide insights into various aspects underpinning the conceptualization and nature of trust – e.g., as the foundation for interpersonal relationships and cooperation, and as the basis for stability in social institutions and markets. However, robust and systematic results into who and why people trust cannot be easily distilled (cf. 'conceptual confusion' [1]. Looking at conceptualizations of trust, among others, psychologists have stressed the role of personality, economists highlighted rational choice, and sociologists have focused on social structures. Consistent trust typologies accompanied by a set of trust constructs are, therefore, not evident.

T. Tryfonas and I. Askoxylakis (Eds.): HAS 2015, LNCS 9190, pp. 271–283, 2015.
DOI: 10.1007/978-3-319-20376-8_25

Moreover, investigation into trust aspects in the context of Information and Communication Technologies (ICTs) which are considered to be an important factor in the adoption of new technical solutions, is arguably, even less consistent. The reason for this is that although trust research has received sufficient attention, studies tend to readily assume that trust is intrinsically beneficial dismissing dependencies such as the context or situation at a given moment in time [2]. Therefore, in order to better understand people and their trust perceptions and appetite towards digital technologies, particularly, Internet-based applications and platforms, this exploratory paper sets out to yield insight into the dynamics of trust and trustworthiness. It will draw particular attention to trust conceptualizations associated with well-placed trust and user attributes.

In its approach, the paper aims to unravel the intricacies of the mechanisms of trust vis-à-vis ICTs and the contextual logic guiding deployment and the user experience, necessitating a view of trust in the digital realm as a dynamic process. Renewed attention is needed to make the networked conditions apparent that underpin user practices, together with a reassessment of the dynamic and open-ended flow of experiences that guides these practices. In other words, the parameters of the 'fabric of trust' are approached as embedded (user-centric and networked-centric) relationships underpinning a usable and secure ICT design process (cf. [3]). The reason for this is that trust in ICTs such as the Internet and Internet-based applications can be seen to erode. While in the early days, trust was one of the drivers that led to a self-reinforcing cycle of (largely beneficial) socio-economic activity facilitating distribution, sharing and collaboration, the situation is different today. Concerns and reduced trust in sensitive data and assets being treated properly can be detected (globally) such as increased criminal activity is affecting more citizens; business models are becoming less transparent, including 'hidden' business roles such as information aggregators and brokers, profilers and networks; an increasing asymmetry of information and control between users and businesses and governments exists; and, privacy is increasingly difficult to maintain, thanks to social networks, super-cookies, location-sensitive services, data aggregation and profiling, and so forth.

In this context, studies have shown that if users trust the ICT system too much (i.e. assume it is more trustworthy than is actually the case), they are exposing themselves to risks and may suffer harm which can also reduce their level of trust in any system in the future. If users trust the system too little (i.e. assume it is less trustworthy than is actually the case), they are failing to benefit from using the system in high-value applications. While such an imbalance - between the level of trust and the level of trustworthiness of services and applications – needs to be tackled and, at minimum, reduced usability issues (that underpin trust related decisions) tend to be overlooked or do not go easily hand-in-hand with trust attributes [2].

Against this backdrop, we give a high-level description for analyzing and understanding the current/experienced trust level of individual users towards online ICT systems and results from a small-scale experiment in the context of Distributed Attack Detection and Visualization (DADV). Our aim is to capture the aforementioned aspects and to reflect on appearing trust tendencies. In doing so, our study adopts a two-step approach deploying survey research to deliver evidence-based conclusions.

2 In ICTs We Trust…

ICTs such as the Internet are said to have historically coevolved with the public who uses them, as well as with the larger economy of inscription. Put aptly by [4], who has provided ample evidence about media and ICT more broadly, they can be defined "as socially realized structures of communication, where structures include both technological forms and their associated protocols, and where communication is a cultural practice, a ritualized collocation of different people on the same mental map, sharing or engaged with popular ontologies of representation" (2006:7). In its ability to connect people across time and space, the power of online (often referred to by Web 2.0 [5]) is rooted in facilitating a range of easily accessible and scalable channels through which interactions can occur. It includes systems that support one-to-one, one-to-many, and many-to-many interactions. Many of these kinds of interactions opened up a myriad of new possibilities for online connections, supporting the generation of 'digital spaces' for people to gather, participate and create, and users to form (e.g. performative innovation, networked publics).

Designing for trust in mediated spaces and interactions has, therefore, become under renewed scrutiny. Understanding the development, maintenance, and enhancement of online trust is of great importance to those involved in the successful design and implementation of digital applications and services. The reason for this is the general belief that trust is central to adoption [6]. Much attention has been given to related elements such as maximizing perceived trustworthiness in e-government services [7] and e-commerce [8], trust cultures [9], and communicating trustworthiness in designing Web sites [10].

In the huge volume of trust research that is available, we have sought to focus on several studies that recognized the need for models of trust and credibility in technology-mediated interactions, particularly, those that aimed to be not-domain specific and technology-independent [1, 2, 7]. These models can be said to offer guidance for researchers across disciplines examining a range of technologies and contexts, thereby highlighting multiple subcomponents, arguably, associated with antecedents (i.e. preconditions of trust), processes of trust building (e.g., interdependence), the context of shaping trust-building (e.g., social relations, regulation), decision-making processes in trust (e.g., rational choice, routine, habitual), implications and uses of trust (e.g., interpersonal entrepreneurial relations, moralistic trust), and lack of trust, distrust, mistrust and repair (e.g., risks, over-trust, trust violations) [11, 12, 14].

What trust models tend to have in common are a categorization by trust referent, and which typically tend to be the characteristics of the trustee (e.g. morality, caring, honesty, willingness to be vulnerable to another). In other words, such models have tended to highlight 'well-placed trust' in their focus on drawing out (sub)components of (perceived) trustworthiness as attributes of the trusted system or party from the trustor's stance, while less attention has been given to the trustworthy attributes, or behavior of the trusted actor [2, 13]. Therefore, we have sought to explore this linkage between ICTs and different trust-related user experiences guided by different sets of trustor attributes.

3 Two-Step Approach

In order to explore the conceptual dynamics underpinning trust-related user experiences and sets of trustor attributes, a two-step approach was deployed. The first step consisted of survey and interview research where respondents were derived from members of the public, the business community, and governmental institutions (February and March 2013, n = 203). Based on a thorough literature review focusing on generic trust models in the design of ICTs supporting (mediated) transactions, the exploratory survey was developed to draw out several key aspects associated with trust in this context. In particular, antecedents, processes of trust building, the context of shaping trust-building, decision-making processes in trust, implications and uses of trust, and lack of trust, distrust, mistrust and repair [1, 2, 7, 11, 15].[1] The purpose was to learn about the combined underpinnings of relevant trust drivers independent from specific technologies and domains. While the first step served mainly to learn about combined constructs in trust-related experiences and attributes, the second step was to conduct a 'segment-specific' analysis so as to learn about different types of subjective trust-related user experiences and attributes in this context. Examining the results of the (end user) survey (n = 90) linkages between different sets of trustor attributes could be associated with trust-related concepts of (1) *Trust stance*: the tendency of people to trust other people across a wide range of situations and persons; (2) *Trust beliefs* in general professionals; (3) *Institution-based trust*; (4) *General trust* sense levels in online applications and services; (5) *ICT-domain specific* sense of trust levels; (6) *Trust-related seeking behavior*; (7) *Trust-related competences*; and, (8) *Perceived importance of trustworthiness design elements*. And, which underpin the segmentation of trust-related user experiences on trustor attributes, thereby, arguably, supporting the estimation of a user's trust level of the ICT system.

4 Results

4.1 Analyzing Trust Levels

It is our aim to assess the relative importance of the trust-related concepts from the literature towards predicting the actual trust in (a set of) online technologies/services, that is, online stores, social networks, professional online networks, online governmental services, online banking, online health services, and online review sites. These trust levels should be considered as general and a priori trust levels towards a particular set of technologies/services, hence, not towards any specific application or on the basis of a concrete experience.[2] We performed the following analysis and present briefly their results.

[1] These constructs were operationalized with using five-point rating scales open questions, checklist questions, and ranking questions.

[2] For instance, we asked in general about trust in online stores and not specifically about trust in individual stores such as Amazon. Users without first hand experience could leave the question unanswered; hence they were not forced to express their opinion.

First, the average trust level vis-à-vis several online technologies was explored. Roughly, we can define three clusters of online technologies based on these trust levels: (1) a cluster containing online banking, online governmental services and online stores with high range trust levels and substantial differences in levels of trust between these three technologies; (2) a cluster with midrange levels of trust containing professional online networks and online health services; and, (3) a cluster with somewhat lower trust level containing social network sites and review sites.

Next, pairwise T-tests were conducted to investigate differences in trust level scores between the various online technologies (see Table 1). These results indicate that a few exceptions, notwithstanding average trust levels, do differ significantly when comparing the various online technologies, and which indicates that various set of technologies are not trusted equally.

Table 1. Pairwised t-tests on trust levels

	Mean 1	Mean 2	Std. Dev. 1	Std. Dev. 2	N	t	Sig. (2-tailed)
Online stores - Social networks	2.09	2.68	.690	1.061	127	-6.436	.000***
Online stores - Professional online networks	2.09	2.31	.678	.916	123	-2.575	.011*
Online stores - Online governmental services	2.10	1.88	.671	.816	129	3.009	.003**
Online stores - Online banking	2.08	1.72	.678	.835	130	5.132	.000***
Online stores - Online health services	2.12	2.40	.671	1.137	100	-2.480	.015*
Online stores - Online review sites	2.13	2.75	.667	.734	122	-7.363	.000***
Social networks - Professional online networks	2.69	2.32	1.049	.933	121	4.781	.000***
Social networks - Online governmental services	2.71	1.90	1.074	.828	126	7.996	.000***
Social networks - Online banking	2.70	1.72	1.083	.845	126	9.843	.000***
Social networks - Online health services	2.64	2.38	1.030	1.144	100	1.922	.058
Social networks - Online review sites	2.76	2.76	1.063	.736	119	.081	.936
Professional online networks - Online governmental services	2.32	1.88	.926	.795	123	5.598	.000***
Professional online networks - Online banking	2.34	1.70	.924	.789	122	6.566	.000***
Professional online networks - Online health services	2.35	2.40	.962	1.142	99	-.449	.654
Professional online networks - Online review sites	2.38	2.72	.905	.738	118	-3.660	.000***
Online governmental services - Online banking	1.91	1.73	.818	.837	128	2.234	.027*
Online governmental services - Online health services	1.92	2.38	.837	1.144	100	-4.912	.000***
Online governmental services - Online review sites	1.91	2.74	.793	.736	122	-8.122	.000***
Online banking - Online health services	1.68	2.40	.750	1.137	100	-6.888	.000***
Online banking - Online review sites	1.76	2.75	.844	.734	122	-10.648	.000***
Online health services - Online review sites	2.42	2.73	1.139	.754	98	-2.400	.018*

*= sig.on .05 level, **= sig. on .01 level and ***=sig. on .001 level

Furthermore, in order to investigate the transferability of trust over the various technology domains, a correlation analysis was conducted (see Table 2).

Table 2. Trust levels correlations matrix

Correlations Matrix						
	Online stores	Social networks	Professional	Online governmental	Online banking	Online health
Online stores	1					
Social networks	.364**	1				
Professional online networks	.325**	.649**	1			
Online governmental services	.407**	.308**	.498**	1		
Online banking	.476**	.354**	.239**	.447**	1	
Online health services	.307**	.229*	.445**	.591**	.447**	1
Online review sites	.136	.251**	.263**	-.083	.169	.095
** Correlation is significant at the 0.01 level (2-tailed).						
* Correlation is significant at the 0.05 level (2-tailed).						

The results indicate that all correlations are positive. This suggests that higher trust levels for one type of online technology seem to go hand-in-hand with higher trust levels for other types and which is a prerequisite for trust being transferable. In terms of strength of the correlations, we observe a high number of high strength/significant correlations

between different trust domains. In particular, social and professional network sites do correlate rather high (r = .649) and as such seems to be distinct from online governmental services, online banking, online health services. Online review sites do correlate moderately with social and professional network sites while seemingly being uncorrelated to the other technologies. From these exploratory findings, we may carefully assume that although a couple of technologies seem distinct from others, trust is - to a certain extent - transferable from one particular technology/service to another.

Next, we looked into what trust-related concepts can be predictive towards these trust levels. In order to assess the internal validity of various scales, factor-analyses were conducted for each of the trust concepts measured on a scale level. Based on these results following trust constructs proofed to show sufficient to excellent internal validity to construct scales: Trust stance (2 item scale, α = .79); Structural assurance (3 tem scale, α = .66); Trust related seeking behavior (5 item scale, α = .86); and, Trust related competences (4 item scale, α = .88).

The results highlight that trust levels vary over different technologies/services. For example, significant predictors for online banking services includes trust stance (β = .291), the design elements 'works well technically' (β = .231) and 'displays seals of approval (β = .239) and trust related seeking behavior (β = −.256). The regression model for online banking has on explained variance of 21,2 %. These results indicate the importance for online banking services to work well technically and to display seals of approval. Trust related seeking behavior can also be seen in this context as an indication of low trust. Finally, a high predisposition to trust (trust stance) helps in building trust towards online banking services.

Diverging sets of trust drivers can be identified for each type of digital technology/service, suggesting the need for a more tailored approach when designing trustworthy applications. However, despite these differences some clear patterns and communalities over the different sets of technologies could be identified. For example, the results suggest that for online technologies, such as governmental services and online banking (that emanate from the public sector or from private sectors where there is a rather strong regulation in place and people also have a strong physical interaction with) trust stance,[3] is the driving factor.

If the main interactions with technologies are of a public nature, and tend to occur in the online realm and depending on the goodwill of the other party or other people in the community (such as online stores, social networks and online review sites), structural assurance,[4] seems to be the driving factor. In only a few cases, displaying a seal of approval is crucial to elicit trust (particularly, for online health services). Other design elements like 'works well technically' and 'easiness to use' seem in certain contexts to facilitate eliciting trust. Furthermore, previously experienced harms can impact trust substantially. Interestingly, it is not so much the type of harm (bullying, fraud etc.) as it is the context (for instance, health-related) wherein the harms are experienced that is predictive towards future trust. As trust is to a certain extent transferable towards other

[3] Or, or the general tendency people have trust in others.

[4] Or, the belief that someone thinks that structures such as regulation and safeguards exist that are important to successfully complete an action.

domains (such as from health to online stores), impactful harms experienced in one domain may significantly lower trust levels over several other domains. Health is a domain wherein harms are likely to be experienced as impactful.

4.2 Analyzing User Segments

For segmentation purposes, a K-means clustering was performed and an Anova analysis was conducted to test for each item whether statistical significance differences could be retrieved between the uncovered trust-related user experience segments. Some iterative clustering and testing led us to a four segments solution to best explain differences in trust-related user experiences. These segments can be represented by the following terms: High trust (HT), Ambivalent (A) trust, Highly active trust seeking (HATS) and Medium active trust seeking (MATS). They differ on a number of aspects (see below), however, based on our analyses, three major concepts are sufficient to explain their core differences. The three underpinning concepts are 'trust stance' (e.g., 'I usually trust a person until there is a reason not to'), 'motivation to engage in trust-related seeking behavior' (e.g., 'I look for guarantees regarding confidentiality of the information that I provide') and 'trust-related competences' (e.g., 'I'm able to understand my rights and duties as described by the terms of the application provider'). They could be measured on 3, 7 and 4 item-scale with a reliability coefficient of .69, .89 and .87 respectively (Table 3).

Table 3. Segmentation results for the three underpinning concepts

	Total (n = 90)	HT (n = 24)	HATS (n = 28)	MATS (n = 18)	A (n = 20)	Anova	
	Mean	Mean	Mean	Mean	Mean	F	Sig.
Trust stance	3,22	3,85	3,15	2,86	3,50	7,260	**,000**
Trust related seeking behaviour	3,52	3,14	4,27	3,34	3,01	4,383	**,000**
Trust related competences	2,44	2,71	2,42	2,94	1,44	2,361	**,000**

The user experience for each of the segments can be characterized as follows:

- The "HT" segment shows a high level trust stance. This means an overall high trust level for the various online applications (e.g. social networks and online banking), accompanied by only few trust seeking behaviors (e.g. checking trust seals), even though the competences are present to cognitively assess the trustworthiness of online applications and services;
- The "HATS" segment displays a high level of trust seeking beyond the mere scanning of trustworthiness cues. It also suggests that individuals are informed about procedures in case of harms and misuse. It points to the capacity of certain competence level that facilitate the assessment of trustworthiness and to the possession, at least, of a minimal understanding of the rules and procedures to look for in case of complaints and misuse. Varied trust stance and trust levels could be observed, including medium to low stance/trust levels;

- The "MATS" segment displays similar characteristics as the 'High Active'. The difference is that the trust seeking behavior is not so apparent. Although drivers for trust seeking behavior (e.g. a low trust stance) and competences to assess trustworthiness can be observed, people's motivation may not be absent to look for trustworthiness cues.
- The "A" trust segment seems to highlight a clear perceived inability to assess the trustworthiness of online applications and services and which may be explained by the personal competence level. Hence, only few active trust seeking behaviors can be observed, yet do not equal low trust levels per se. Trust seems to be derived from either the general trust stance or basic heuristics, such as 'public organizations are more trustworthy than commercial companies'. The "Ambivalent" nature of this user experience might be explained by a failure to cognitively assess the trustworthiness and a certain need to trust in order to avoid, or to lower the omnipresence of cautious and other negative feelings, and which is a so-called 'forced trust' (that is, trust without trustworthiness evidence and with a possible presence of cautious feelings). These findings point to understanding trustworthiness indicators based on the experience of others (referrals), as the main source of 'trustworthiness information' that is accessible for this cluster, and underlying the outcome of the trustworthiness assessment.

The user experience for the "HT" segment can be characterized by a high level trust stance. This means an overall high trust level for the various online applications, such as social networks and online banking, accompanied by only few trust seeking behaviors, such as checking trust seals, even though the competences are present to cognitively assess the trustworthiness of online applications and services.

For the "HATS" segment, the user experience can be highlighted in terms of a high level of trust seeking behavior beyond the mere scanning of trustworthiness cues. It also suggests that individuals are informed about procedures in case of harms and misuse. It points to the capacity of certain competence level that facilitate the assessment of trustworthiness and to possess, at least, a minimal understanding of the rules and procedures to look for in case of complaints and misuse. Varied trust stance and trust levels could be observed including medium to low trust stance/trust levels.

For the "MATS" segment, the user experience is similar to the "Highly active" one, yet, here, trust seeking behavior is not so apparent. Thus, while drivers for trust seeking behavior, such as a low trust stance, are present as well as competences to assess trustworthiness, people's motivation may be absent to look for trustworthiness cues.

The "A" trust segment seems to highlight a clear perceived inability to assess the trustworthiness of online applications and services and which may be explained by the personal competence level. Hence, only few active trust seeking behaviors can be observed, yet do not equal low trust levels per se. Trust seems to be derived from either the general trust stance or basic heuristics, such as 'public organizations are more trustworthy than commercial companies'. It seems that the "Ambivalent" nature of this user experience can be explained by a failure to cognitively assess the trustworthiness and a certain need to trust in order to avoid, or to lower the omnipresence of cautious and other negative feelings, and which is a so-called 'forced trust' (that is, trust without trustworthiness evidence and with a possible presence of cautious feelings). These findings point

to understanding trustworthiness indicators based on the experience of others (referrals), as the main source of 'trustworthiness information' that is accessible for this cluster, and underlying the outcome of the trustworthiness assessment.

By segmenting trust user experiences we are able to pinpoint different sets of trust drivers, competences, attitudes and behaviors. This suggests that (computational) trust models may benefit from including trustor's attributes as model parameters. In this view, and as a next step, we can envision that a user without previous first-hand experience with any system, completes a short intake questionnaire before interacting with the system. That intake questionnaire could serve to (1) to assign a particular user to any of the four clusters of trust related user experiences, and (2) to help initiate and, at a later stage, validate the model by calculating initial trust values.

In the next section, therefore, we outline some preliminary findings of a small-scale experiment of a Distributed Attack Detection and Visualization (DADV) system.

4.3 Validating Results

In order to indicate how different sets of trustor's attributes (trust stance, motivation to engage in trust seeking behavior and competences) may relate or impact different model parameters during a trust estimation phase, an experiment was conducted in October 2014. Participants were asked to test and evaluate an online security service of a fictitious provider providing a service, called DADV, to detect virtual attacks on devices connected to the Internet, such as personal computers. The experiment was performed for two versions of the online service; on the one hand, a service where administrators are responsible for detecting and mitigating attacks, on the other hands, an automated service where all tasks are performed by sophisticated tools (Table 4)

Table 4. Segmentation results for experiment participants (n = 27)

	Total (n = 27)	HT (n = 5)	HATS (n = 4)	MATS (n = 10)	A (n = 8)	Anova	
	Mean	Mean	Mean	Mean	Mean	F	Sig.
Trust stance	2,65	3,40	2,63	2,30	2,63	4,519	**,012**
Trust related seeking behaviour	2,16	2,06	2,82	1,89	2,25	6,879	**,002**
Trust related competences	3,53	3,80	4,13	3.58	3.58	3,067	**,048**

In order to assess whether the four segmentation solution could be deployed, additional empirical research was carried out. For this purpose the survey was dispersed using several Living Lab panels in September 2014 (n = 89). The same analytical steps were followed as above. While some minor variations between the two exploratory analyses could be detected, the dominant drivers that seem to characterize users in each segment appear to be relatively constant. Thus, the findings seem to correspond to the previous ones indicating that the three underpinning users' attributes appear as statistically significant difference. More specifically, we observe that the combined aggregate

factor of 'competences' and 'seeking behavior' is again higher for the HATS segment. This finding justifies our approach to correlate higher values of this factor with a more accurate estimation. Furthermore, it is confirmed that a high level of 'trust stance' results to trustworthiness overestimation (misplaced trust) and vice versa (presence of over-cautious users). From those who filled out the intake survey, 27 individuals participated in the second phase of the evaluation (the actual experiment). And which took place on two consecutive days (separate Vanilla and OPTET TTM evaluations), lasting for about 90 min on October 8 and 9, 2014.

In order to validate the trust initialization participants were asked to report their initial trust towards the system before having any other evidence for its performance. To do so, each participant engaged with the DADV system, separately for each version during two different days, starting with the administrated and then with the automated one. After logging in to the online website (and before any attack was performed), they were given the opportunity to access the 'about page' and familiarize themselves with the activated version. This webpage provided general information of the system functionality and a high-level description of its expected trustworthiness. Furthermore, users who had noticed and clicked on a distinguishable hyperlink were redirected to a more detailed webpage, which explicitly mentioned each system's actual trustworthiness in terms of the metric under interest. In this way we could validate the effects of 'seeking motivation' on the initial trust level of each segment (Figs. 1 and 2).

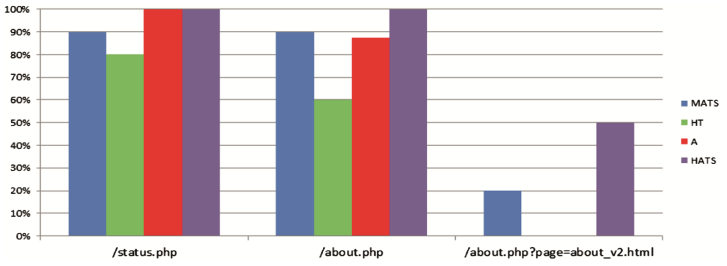

Fig. 1. Percentage of users in cluster that visited the webpages of the administrated DADV at least once. MATS depicts medium-active trust seeking, HT depicts High Trust, A depicts Ambivalent, and HATS depicts highly-active trust seeking.

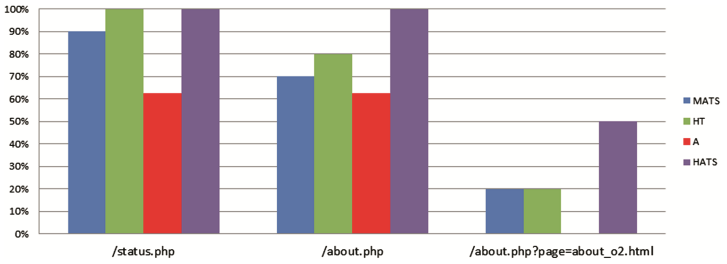

Fig. 2. Percentage of users in cluster that visited the webpages of the automated-enabled DADV at least once.

Afterwards, they observed the service performance for a sequence of 10 attacks that were identical for both DADV cases. During each attack, they could navigate to the 'health statistics page', which was providing a holistic view of the system status. At the end of each attack a message was appearing indicating whether the provider succeeded in preventing any network host from being attacked, or not. These pop-up messages also contained a link to a questionnaire where users were asked to indicate their current trust level that the provider would prevent future attacks from compromised honeypots to their computers.

We observed that the segment HATS achieved the highest percentages in all six webpages of interest (we excluded the Home page because it has been loading and was periodically refreshed automatically). Furthermore, the 'A' segment achieved the lowest percentage in 5 out of 6 webpages. The other two segments had different behavior during the two days.

Table 5 depicts the trust level of the users after the first day of the experiment, where the administrated DADV succeeded in protecting the users' infrastructure 6 out of 10 times. We observed that some members of the HT segment were the only ones to feel extremely secure. Note that the HT mean value for the Trust stance concept in is the highest among all segments. In general, most participants of the other segments were either moderately or very secure.

Table 5. Answers to the Question "To what extent did you feel protected using the DADV service?" (Day 1) (N = 27).

	Slightly	Moderatedly	Very much	Extremely	Total
MATS	7.4 %	14.8 %	14.8 %	0 %	37 %
HT	0 %	3.7 %	11,1 %	3.7 %	18.5 %
A	3.7 %	18.5 %	7.4 %	0 %	29.6 %
HATS	0 %	7.4 %	7.4 %	0 %	14.8 %
Total	11 %	44 %	41 %	4 %	100 %

Table 6 presents the effect on users' trust of having more timely information about on-going attacks. Note that the only segment whose members would not prefer to receive such notifications is the 'A' one, which has the lowest mean value for the Trust-related competences concept. In general, most participants seem to perceive such a feature positively, in particular the Highly-active trust seeking ones who have the highest mean value for the Trust-related competences concept.

The participants were also asked to indicate the extension to which they felt protected using the DADV service. Note that for none of the following graphs significance tests could be performed as most cells had expected count less than 5, and which, as stated in D2.3 makes a large scale experiment rather valuable. Also, when asking the respondents about a possible change in trust level vis-à-vis receiving real-time alerts about attacks, some 45 % of the respondents indicated that it may likely change (m = 3.78, SE = .154, SD = .801). Lastly, the findings have shown that the respondents (N = 24)

seemed to prefer the automated-enabled version (8.3 % preferred Day 1 versus 91.7 % that preferred Day 2, m = 1.92, SE = .058, SD = .282). When they were asked to justify their choice, most participants stated that automated-enabled version managed in preventing more compromised sensors from attacking their infrastructure than the administrated one.

Table 6. Answers to the Question "Would your level of trust in the service to protect your infrastructure change if you were to receive real-time alerts of attacks?" (Day 1) (N = 27).

	Slightly	Moderatedly	Very much	Extremely	Total
MATS	0 %	7.4 %	25,9 %	3.7 %	37 %
HT	0 %	14.8 %	3.7 %	0 %	18.5 %
A	3.7 %	7.4 %	7.4 %	11.1 %	29.6 %
HATS	0 %	3.7 %	7.4 %	3.7 %	14.8 %
Total	3.7 %	33.3 %	44.4 %	18.5 %	100 %

Finally, we noticed that a very accurate initial trust value for the HT and HATS segments could be estimated, while the relative error for the A and MATS segments is 10 % and 20 %, respectively. With regard to the trust dynamics, the (computational) models are aligned with the expected user reactions for most segments; namely trust should not decrease after a success and should not increase after a failure. The only exception is for the 'MATS' segment, which appears to constantly increase with the number of trials. This can be attributed to the error in estimating that particular initial value; in such cases the system may result in negative values for one or both update coefficients. We also observe a close estimate of the average trust of the HATS and Ambivalent segments, while for the rest segments the relative error is less than 15 %.

5 Conclusions

In this paper the conceptual dynamics underpinning trust-related user experiences and sets of trustor attributes have been explored. From the segmentation research presented, we conclude that our analysis seems valid and results to the capacity to steadily detect dominant drivers that affect the subjective nature of trust. Based on these findings we derived the expected users behavior, considering also the technical factors that determine system performance. To this end, trust was explicitly formulated as a function of both aforementioned aspects, while shaping the expected behaviors of each segment. However, we did not always manage to closely estimate the actual trust values. We observed that only a small subset of them is adequate so that to estimate trust with great accuracy, while also shaping the actual user's reactions.

Concerning our future work, we aim to perform a large scale experiment with respect to the number and profile of participants and the number of trials observed from each

individual. We also aim to validate the user segmentation not only in terms of comparison with other related studies (as the methodology followed here), but also based on their actual attributes and behavior (e.g., pages visited, duration of visit). The reason for this is that there is always the possibility of users not being truthful when answering the questionnaire or competent enough to understand the questions, thereby highlighting and reflecting findings concerning the lifecycle of trust.

References

1. McKnight, H., Chervany, N.: Trust and distrust definitions: one bite at a time. In: Falcone, R., Singh, M., Tan, Y.H. (eds.) Trust in Cyber-Societies: Integrating the Human and Artificial Perspectives, pp. 27–54. Springer, Berlin (2001)
2. Riegelsberger, J., Sasse, M.A., McCarthy, J.D.: The mechanics of trust: a framework for research and design. Int. J. Hum. Comput. Stud. **62**(3), 381–422 (2005)
3. van der Graaf, S.: Imaginaries of ownership; the logic of participation in the moral economy of 3D software design. Telematics Inform. Spec. Issue Ethics Inf. Soc. **32**(2), 400–408 (2015). Elsevier
4. Gitelman, L.: Always Already New: Media, History and the Data of Culture. MIT Press, Cambridge (2006)
5. O'Reilly, T.: What is web 2.0. http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html (2005). Accessed 4 Sept 2014
6. Lacohée, H., Cofta, P., Phippen, A., Furnell, S.: Understanding Public Perceptions: Trust and Engagement in ICT-Mediated Services. International Engineering Consortium, Chicago (2008)
7. Tan, Y.-H., Thoen, W.: Toward a generic model of trust for electronic commerce. Int. J. Electron. Commer. **5**(2), 61–74 (2001)
8. Cheshire, C.: Online trust, trustworthiness, or assurance? Dædalus J. Am. Acad. Arts Sci. **4**, 49–58 (2011)
9. Karvonen, K., Cardholm, L., Karlsson, S.: Cultures of trust: a cross-cultural study on the formation of trust in an electronic environment. In: Proceedings of the Fifth Nordic Workshop on Secure IT Systems, NordSec (2000)
10. Nielsen, J. (n.d.): Trust or bust: communicating trustworthiness in web design. http://www.nngroup.com/articles/trust-or-bust-communicating-trustworthiness-in-web-design/. Accessed 15 Feb 2015
11. Lyon, F., Möllering, G., Saunders, M.N.K. (eds.): Handbook of Research Methods on Trust. Edward Elgar, Cheltenham (2012)
12. Li, F., Kowski, D.P., van Moorsel, A., Smith, C.: Holistic framework for trust in online transactions. Int. J. Manag. Rev. **14**, 85–103 (2012)
13. Möllering, G.: Trust: Reason, Routine, Reflexivity. Elsevier Ltd., Oxford (2006)
14. Sztompka, P.: Trust: A Sociological Theory. Cambridge University Press, Cambridge (1999)
15. Barbalet, J.N.: A characterization of trust, and its consequences. Theor. Soc. **38**(4), 367–382 (2009)