Mental Models: An Approach to Identify Privacy Concern and Behavior

Kovila P.L. Coopamootoo
School of Computing Science
Newcastle University
Newcastle Upon Tyne, NE1 7RU, UK
kovila.coopamootoo@newcastle.ac.uk

Thomas Groß
School of Computing Science
Newcastle University
Newcastle Upon Tyne, NE1 7RU, UK
thomas.gross@newcastle.ac.uk

ABSTRACT

Elicitation of privacy concern is a difficult task due to its sensitivity to priming effects. The methodology poses the risk of triggering cognitive processing and activation of mental models leading to responses that might not usually arise in everyday interactions. Privacy is highly contextual and end users' privacy requirements vary with time, beliefs, experience and perceived social norms. This position paper proposes mental models theory as a paradigm for identifying, measuring and modeling end user privacy attitude, concern, intention and behavior. We review background literature and make a case for how conceptual representation of end user mental models of online privacy provide a valuable structure to investigate links between privacy concerns and behavior. It does so by enabling depiction of conceptual relationships that hold semantic information, which could portray end users' cognitive processing and reasoning such as identifying concepts that are strong for different users or lead to most associations. Thus mental models would provide a framework for segmentation and identification of pathways from concerns to behaviors.

1. INTRODUCTION

Although end users report privacy concerns, their behavior often do not match these concerns [35, 1]. Privacy concerns can be thought to represent an instantiation of privacy attitude since it refers to a certain tendency or disposition. The incongruence between attitude and behavior is not a new phenomenon. LaPierre refuted the belief that behavior naturally follows attitude with his study of whether restaurants and hotels would serve a Chinese couple [26]. He found that only one out of 251 establishments visited actually refused to serve the visitors although only 10% indicated a willingness to serve them in a previous questionnaire. Since then cognitive and social psychologists have investigated the discrepancy and proposed a variety of theories related to motivation and volition [2], self-regulation [24, 5] and problem solving and performance control [22].

Symposium on Usable Privacy and Security (SOUPS) 2014, July 9–11, 2014, Menlo Park, CA.

In addition privacy presents a supplementary set of problems due to its tacit and highly contextual nature [28]. To address the problem of predicting privacy behavior from attitudes and concerns, we first propose investigation of the links between attitudes, concerns, behavioral intentions and behavior. This involves inquiry of end users' motivational states and self-regulation abilities that would favour certain behaviors, of mechanisms that support formation and enactment of intentions and of factors that influence the transition from cognition to action. Second such investigation has to account for psychological biases that underpin judgment and decision making under uncertainty and distinguish between dual-process modes of thinking.

We believe mental models will support the two above propositions and the conceptual representation of mental models will provide a rich foundation to understand human thinking and reasoning [21]. In this paper, we propose mental models theory as a paradigm to generate conceptual representations of end users' mental models of privacy online. We present a first study that used Amazon's Mechanical Turk and cognitive mapping technique to elicit and depict mental models. Through structural analysis we show that the cognitive maps could support segmentation of user concerns.

In the next sections we review research on the attitude-behavior link before introducing mental models and cognitive maps. We discuss our proposition together with brief preliminary findings of our study before concluding with future works.

2. PRIVACY ATTITUDE AND BEHAVIOR

Although various definitions of attitudes have been proposed, Perloff proposes a version that encloses the essential points of the concept as:

DEFINITION 1 (ATTITUDE [31]). *An attitude is defined as a learned, global evaluation of an object (person, place or issue) that influences thought and action.*

First people are not born with attitudes; rather attitudes are learnt over the course of life through socialisation [31]. Attitudes are developed through encounters with social objects. Second evaluation of encounters can be expressed through thoughts, feelings, intentions to behave and behavior. Third, attitudes help people to organise their social world and hence influence behavior [31]. Privacy attitude in literature is often linked to privacy concern such as Westin's Privacy Segmentation Index [25].

Social psychology distinguishes between the availability of an attitude and its accessibility [16]. While availability of an attitude refers to whether it is present within an individual's cognition, accessibility of attitude refers to the ease with which the attitude may be retrieved from memory [14]. Accessibility of an attitude is dependent on the properties of the attitude such as strength of the attitude and aspects of the context that highlight particular attitudes as being relevant [3]. Thus accessible attitudes lead to corresponding behavior more closely than less accessible attitudes [15]. Therefore, whilst individuals have privacy attitudes, these may not be accessible when they interact online. The low accessibility of privacy attitudes during interactions may be first due to the lack of relevance of the online disclosure context to privacy attitudes and second due to the weak strength of the privacy attitude. Rather than extremity in terms of positive or negative value of the attitude, the strength of a particular attitude refers to a consistent, well-rehearsed link between an attitude object and its evaluation [3]. It was shown by Fazio et al. [13] that when the association between an object and its evaluation is strong enough, simply noticing the object would cause evaluation.

Moreover, although attitudes may be available, they may not be active; that is they may not be associated with an object or issue to cause its evaluation [14]. This characteristic of attitudes can also explain the dichotomy: although individuals hold privacy attitudes, they may not be able to associate these attitudes with their online activities.

Attitudes can have multiple cognitive and affective components thus leading to different evaluations of an object. These ambivalent attitudes are unstable since the evaluation expressed at a particular moment depends on the elements of the attitude that is most accessible at that time [11]. Privacy attitudes for the online environment can be said to be ambivalent since a variety of different disclosure contexts exists online that makes it hard for end users to maintain stable and strong evaluations. This property of online privacy attitudes may also explain why end users do not behave according to their privacy attitudes.

Furthermore an attitude towards an object is viewed as related to the person's intentions to perform a variety of behaviors with respect to that object. The relation is between attitude and the set of intentions as a whole and the attitude toward an object might not be related to any specific intention with respect to the object. Each intention is viewed as being related to the corresponding behavior via the target object at which the behavior is directed, the situation and the times of execution. Each of these elements varies along a dimension of specificity [11].

3. MEASURING PRIVACY ATTITUDE AND BEHAVIOR

Measurements of privacy attitudes in previous research have compared measures of broad privacy concerns with self-reports of privacy-preserving strategies, self-reports of past disclosures [1] or with observed disclosure actions [35]. The discrepancy in privacy behavior and attitude may therefore either be due to inconsistency in terms of the specificity of attitude, intentions and behavior or due to taking disclosure actions to be a direct measure of privacy behavior.

Since LaPierre questioned the relationship between attitude and behavior in 1934 [26], there has been a series of discussions, among which calls for better measurement of attitude and behavior such as Schuman and Johnson's call to consider conceptual congruence [17]. Also, Ajzen and Fishbein [18] found measurement issues relating to two types of attitudes: general attitudes towards an object versus attitudes towards performing behaviors with respect to an object or target, that is behavioral intentions. They identified two types of inconsistency: evaluative inconsistency and literal inconsistency.

Evaluative inconsistency occurs when broad attitudes are compared with single behaviors. This is a problem because it is only under certain conditions and/or for certain individuals that general attitudes have strong impact on behavior. Therefore, Westin's Privacy Segmentation Index [25] would not be compatible with specific online privacy behavior.

Literal inconsistency is the inconsistency between behavioral intentions and actions. Such inconsistency can be caused if behavior is not easy to perform as argued by Campbell [7] or if there is a time interval between measurement of intention and assessment of behavior. Also if intentions change during that time, the intentions will tend to be poor predictors of behavior.

4. MENTAL MODELS

Mental models are internalised, mental representations of a device or idea that facilitates reasoning [20]. They are simplistic and small-scale representations of reality [9]. Mental models are valuable because they are the lenses through which individuals see and interact with the world. The lens shapes how individuals interpret the world. Thus by conjecture, mental models would comprise our attitudes, beliefs, opinions, theories, perceptions, mental maps of how things are or should be and frames of reference.

Mental models vary with user expertise and experience. Experts' mental models are richer and more abstract than those of novices. Novices' models represent more concrete levels of knowledge and have a more naive problem representation as they present objects in real time [27]. Moreover experts might aggregate different contexts as having some equivalence while novices would consider them distinctly.

It is also thought that users build and use models to guide the way they learn and interact with computers. Mental models enable users to predict and explain the operation of a target system through internal representations of themselves and the objects they interact with [29]. By interacting with systems, users formulate mental models of the system that need not be technically accurate but are functional that is the model can be 'run' and works within a certain scenario. Since users improve their models with experience, mental models are often incomplete and partial descriptions of the operations of the system.

Furthermore, mental models have been associated with privacy and security research before through analogies and metaphors. These include 'situational faces' [34], 'audience-view' [33], card-based metaphors [38] or modeling of security risks [6] as reviewed in [8].

4.1 Cognitive Maps

Cognitive maps can be regarded as expressions of mental models and cognitive mapping to the task of mapping a person's thinking about a problem or issue. It is a technique used to structure, analyse and make sense of accounts of problems that can be verbal or written. Cognitive map

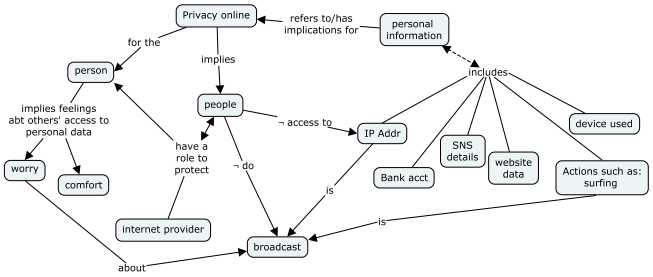

Figure 1: Cognitive Map e.g. for Q1

has had a long history, the idea originally coined to depict mental representations of the routes and paths of the environment used by people and rats [36]. However Axelrod [4] used it as a 'map of cognition' while Eden later used it to refer to a map 'to aid cognition' [12]. Axelrod's map of cognition has been used in artificial intelligence [23] and experimental research such as system dynamics [10]. In our research, we also use cognitive maps as originally referred to by Axelrod. However the 'mental model uncertainty principle', that is mental models are not directly accessible or observable, accentuates the inherent problem of representing mental models [32]. Also an agreed upon cognitive mapping methodology is not yet available between research domains [37].

5. OUR APPROACH

In this section, we present a study aimed at eliciting and developing user privacy mental models. We present our design followed by an analysis of the structural properties of the models.

5.1 Design

Our on-going research includes elicitation of privacy mental models via Amazon's Mechanical Turk. We are conducting between-subject studies involving different framing of privacy related questions requiring 100 to 250 words responses. Our first question **Q1** was 'What does privacy online mean to you?'. Having noticed concrete stories and low-level actions, we were intrigued and wondered if the word 'privacy' would even arise in the different set of questions **Q2**: 'What do you usually use the internet for?' and 'What is important to you when you are online?'. We have recently launched another version with **Q3** 'What does sharing online mean to you?'. As a test case for our methodology we collected and analysed the data of five partcipants for each of our first two questions.

5.2 Structural Analysis

We used CMapTools to develop cognitive maps [30]. The process involves separating sentences into *distinct phrases* of no more than 10–12 words long (possibly much shorter) and building up a hierarchy of concepts linking to each other while identifying directions.

Definition 2 (Cognitive Map). *A* cognitive map *is a directed, possibly cyclic, vertex-labeled and edge-typed/-labeled multi-graph. The vertices are labeled with distinct concepts. The arrows depict thought processes for a person with links or associations from one concept, the source to another, the sink. An arrow is derived from a one-to-one mapping of a phrase to concept relation. The directed associations could encode cause/effect or means/ends but are not limited to these.*

We first look at the shape of the maps. The different questions give different structures:

- maps for **Q1** have a hierarchical structure pointing towards/from the main concept 'privacy online' and often linking to three clear subordinate but important concepts: the person, personal information or data and other people who can be authorised or not. These link to concrete examples making a three-level graph on average as shown by Figure 1.
- for **Q2**, three maps had a shallow hierarchy leading to the superordinate concept 'person' from information or type of activities, often also leading to the concept of 'friends' or social connections. Therefore the maps show the different activities for which the person uses the online environment. Each of these three maps has one to three longer links that show who the person share specific information with and the benefits of obtaining information on the internet.

Second we identify source and sink vertices. Sink vertices are concepts represented by vertices that have an out-degree of zero. Such concepts have in-arrows only, with no out-arrows. Source vertices are concepts represented by vertices that have an in-degree of zero. Table 1 summarises the sink and source vertices for **Q1** while Table 2 those that emerged from **Q2**.

Third we look at the degrees of vertices which refers to the number of direct links (both input and output). Table 3 and 4 provide the list of concepts that received at least a degree of 3 for each partcipant of **Q1** and **Q2**.

6. DISCUSSION

Cognitive maps as a representation of mental models would be a valuable tool for validation of existing segments or for elicitation of segments. Sink and source vertices together with the degree of vertices point to important concepts or clusters for participants. While sink vertices might potentially help distinguish between varying privacy concerns, source vertices could identify triggers that prime a certain behavior or activate more elaborate mental models of privacy. Concepts leading to multiple sink vertices might indicate their strength. Reachability of concepts and cycles might give further indication of the users' thought processes. It would also be interesting to find out whether vertices trigger links with inhibitory, aversive or approving decisions and behavior. For instance it appears from Table 1 and Table 2 that we are able to identify sink and source vertices. Table 3 shows the high importance of 'personal information' or 'data', the 'person', other 'people' and 'privacy online', Table 4 shows the prominence of the concept 'person' and much of the social benefits of the online environment through 'friends', 'social reach' and 'shopping'.

Table 1: Sink and Source vertices for Q1 for participants 1 to 5

P	Map Sink	Map Source
1	broadcast, comfort	internet provider, IP address, bank account, SNS details, website data, surfing actions, device used
2	person, life	online life, job application
3	inviolable right, people	authorised, SNS protector, account, files, pictures, messages, friends
4	data, identity anonymity history	email, personal identifiers, posts, prying eyes
5	people	criminals, comment to an article or blog, anonymous

Table 2: Sink and Source vertices for Q2 for participant 1 to 5

P	Map Sink	Map Source
1	foreign friends, favourable, trustworthy, careful, bank information	(no tail nodes)
2	projects, knowledge, live chat and video	social websites, film, email, job application, online shop, people
3	better, perfect, high speed	internet connection, travel tickets, bills, to know something, shopping, online work, films, favourite site
4	hackers, banking sites, tv/movies, social media	quickly, personal financial info, news, emails, others
5	traffic load, power to connect, great	everything, critical factor, undeniable, topics

However, the high degree for participant 1 of **Q2** include 'unknown' and 'known'. This corroborates with Table 2, where the same participant produced sink vertices including 'favourable', 'trustworthy', 'careful' and 'bank information'. Participant 4 has less risk related concepts but mention 'sensitive information' without being prompted about privacy and Table 2 identifies 'hackers', 'banking sites' among sink vertices.

The shape of the graphs together with the lengths of arguments can be an indication of the participants' cognitive processes with respect to the question but it can also be influenced by thoughts and ideas that are more salient at the time of participating in the study. Moreover the shape can be induced by the type of questionning. For instance **Q1** included 'What does . . . mean to you?' whereas **Q2** was 'What is. . . '. This might contribute to **Q2**'s generally shallow map associated with activities.

However further analysis and evidence are required to corroborate these findings across types of maps and establishing whether a particular user belongs to a segment would depend on the consistency of the maps. Also given the 'mental model uncertainty principle' described by Richardson et al. [32], the stability of mental models is questionnable. Thus techniques to ascertain the stability of mental models over time and circumstance akin to those in trait theory [19] would be valuable for the research. In addition our cognitive mapping methdology has not been validated yet nor have we assessed whether other methods would be a more suitable and reliable. Our research agenda includes developing a rigorous, systematic and reproducible methodology and conducting the study with a larger sample.

7. CONCLUSIONS

In this short paper, we make a case for the use of mental models to elicit end users' privacy concerns and behaviors. We presented the characteristics of attitudes and advanced how these can be linked with inconsistent measurements. We introduced mental models as providing a framework that embraces the complex relationship between privacy concern and behavior. We provided an excerpt of our methodology showing the use of cognitive mapping. Our mental model approach first promises content that illustrates intervening links such as behavioral intentions, modes of thinking and reasoning. Second, the structure of mental models contributes towards segmentation and filtering while its semantics present valuable insights in the area of usable privacy and privacy decision-making.

8. FUTURE WORK

While we are still in the initial phases, we aim to:

- validate our elicitation approach by for example first testing different framing of privacy questions across different user groups, second comparing the raw data from Mehanical Turk to those of interviews;
- evaluate our phrase extraction approach such as tapping into linguistic tools;
- differentiate between types of arrows or relationships such as ontology ('is a', 'includes'), constraint (restriction of the application of the concept), cause-effect or negation thereof;
- develop intermediate steps between depiction and analysis to identify and add semantic meaning into the maps;
- conduct further analysis including co-occurence matrix, cognitive distance, cluster analysis, multidimensional scaling and hierarchical cluster analysis;
- develop a composite conceptual map.

Table 3: Degree of vertices for Q1

P	Concepts (degree)
1	personal information (7), person (4), broadcast (4), IP address (3), privacy online (3)
2	data (6), privacy online (4), person (4), unauthorised people (3), stolen (3)
3	data (10), people (5), person (3), inviolable right (3), privacy online (3), unauthorised (3)
4	data (5), person (4), website and service (4), privacy online (3), people (3)
5	information (6), person (5), a right (3), website (3), people (3)

Table 4: Degree of vertices Q2

P	Concepts (degree)
1	person (11), unknown (9), known (5), personal information (4), people (4), website online (3), somebody (3), money (3)
2	person (11), online (6), friends (6), information (4), other (3), profile (3)
3	person (12), friends (4), internet connection (3), SNS (3)
4	person (9), sensitive information (4), shopping (3), TV and movies (3), MTurk (3)
5	person (6), social reach (5), people (3)

Our research plans as elaborated in [8] also include investigation of reasoning within mental models such as inductive or deductive modes and the dual-process models.

Acknowledgments

The authors' research on usable privacy is currently supported by the EU FP7 FutureID project (http://futureid.eu) under GA n^o 318424.

9. REFERENCES

[1] A. Acquisti and J. Grossklags. Privacy and rationality in individual decision-making. In *IEEE Security and Privacy*, volume 3 of *1*, pages 26–33, 2005.

[2] I. Ajzen. *Action Control: from cognition to behavior*, chapter From intentions to actions: a theory of planned behavior. Springer Verlag, 1985.

[3] M. Augoustinos, I. Walker, and N. Donaghue. *Social Cognition: An integrated introduction*, chapter Attitudes. Sage Publications, 2006.

[4] R. Axelrod. *The structure of decision*. Princeton University Press, N.J., 1976.

[5] J. Beckmann and M. Irle. *Action Control: from cognition to behavior*, chapter Dissonance and action control. Springer Verlag, 1985.

[6] L. Camp. Mental models of privacy and security. In *IEEE Technology and society magazine*, volume 28, 2009.

[7] D. Campbell. *Psychology: a study of a science*, chapter Social attitudes and other acquired behavioral dispositions. McGraw - Hill, 1963.

[8] K. P. L. Coopamootoo and T. Groß. Mental models for usable privacy: A position paper. In T.Tryfonas and I.Askoxylakis, editors, *HAS 2014*, volume 8533 of *LNCS*, pages 410–421. Springer Int, 2014.

[9] K. Craik. *The nature of explanation*. Cambridge University Press, 1943.

[10] J. K. Doyle and D. N. Ford. Mental models concepts for system dynamics research. *System dynamics review*, 14(1):3–29, 1998.

[11] A. Eagly and S. Chaiken. *The psychology of attitudes*. Wadsworth Publishing, 1993.

[12] C. Eden. On the nature of cognitive maps. *Journal of Management Studies*, 29(3), 1992.

[13] R. Fazio. *Handbook of motivation and cognition: Foundations of social behavior*, chapter How do attitudes guide behavior? Guildhall press, 1986.

[14] R. Fazio. *Attitude structure and function*, chapter On the power and functionality of attitudes: the role of attitude accessibility. Lawrence Erlbaum Associates, 1989.

[15] R. Fazio and C. Williams. Attitude accessibility as a moderator of the attitude-perception and attitude-behavior relations: an investigation of the 1984 presidential election. *Personality and social psychology*, 51:504–514, 1986.

[16] E. Higgins. *Social psychology: handbook of basic principles*. Guildhall press, 1996.

[17] H.Schuman and M. Johnson. Attitudes and behaviors. *Annual review of psychology*, 2:161–207, 1976.

[18] I.Ajzen and M.Fishbein. Attitude-behavior relations: a theoretical analysis and review of empirical research. *Psychological bulletin*, 84:888–918, 1977.

[19] O. P. John and S. Srivastava. The big five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, 2(1999):102–138, 1999.

[20] P. Johnson-Laird. *Mental models: towards a cognitive science of language, inference and consciousness*. Cambridge University Press, 1983.

[21] P. Johnson-Laird. *How we reason*. Oxford university press, 2006.

[22] R. H. Kluwe and G. Friedrichsen. *Action Control: from cognition to behavior*. Springer Verlag, 1985.

[23] B. Kosko. Fuzzy cognitive maps. *International journal of man-machine studies*, 24(1):65–75, 1986.

[24] J. Kuhl. *Action Control: from cognition to behavior*, chapter Volitional mediators of cognition-behavior consistency; self-regulatory processes and action versus state orientation. Springer Verlag, 1985.

[25] P. Kumaraguru and L. Cranor. Privacy indexes: A survey of westin's studies. Technical report, CMU-ISRI Technical Report, 2005.

[26] R. LaPierre. Attitudes vs. actions. *Social Forces*, 13:230–237, 1934.

[27] J. Larkin. *Mental models*, chapter Expert representations of physics problems. Lawrence Erlbaum Associatesn, 1983.

[28] H. Nissenbaum. Privacy as contextual integrity. *Washington Law Review*, 1:129–136, 2004.

[29] D. Norman. *Human Computer Interaction: some observations on mental models*. Morgan Kaufman Publishers Inc, CA., 1983.

[30] I. of human and machine cognition. Cmaptools.

[31] R. Perloff. *The dynamics of persuasion:communication and attitudes in the 21st century*, chapter Attitudes: definition and structure. Routledge, 2010.

[32] G. P. Richardson, D. F. Andersen, T. A. Maxwell, and T. R. Stewart. Foundations of mental model research. In *Proceedings of the 1994 International System Dynamics Conference*, pages 181–192, 1994.

[33] H. Richter-Lipford, A. Besmer, and J. Watson. Understanding privacy settings in facebook with an audience view. In *1st conf. on usable psychology and security*, pages 1–8, 2008.

[34] S.Lederer, J.Mankoff, and A.K.Dey. Who wants to know what when? privacy preferences determinants in ubiquitous computing. In *CHI '03*, 2003.

[35] S. Spiekermann, J. Grossklags, and B. Berendt. E-privacy in 2nd generation e-commerce: privacy preferences versus actual behavor. In *3rd ACM E-com conf.*, pages 38–47, 2001.

[36] E. C. Tolman. Cognitive maps in rats and men. *Psychological review*, 55(4):189, 1948.

[37] J. A. Vennix. *Mental models and computer models: Design and evaluation of a computer-based learning environment for policy-making*. 1990.

[38] E. Wastlund, J. Angulo, and S. Fischer-Hubner. Evoking comprehensive mental models of anonymous credentials. In LNCS, editor, *Open problems in network security*, volume 7039, pages 1–14, 2012.