Mental Models of Online Privacy: Structural Properties with Cognitive Maps

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

Individuals usually build small-scale representation of reality to help them navigate their environment. Although mental models have been used in HCI before, they mostly occur as analogies and metaphor within the privacy and security research space. The meaning for users, the values associated and reasoning over online privacy have not been investigated before. In our research we explore and depict users' mental models of online privacy through the content, properties and structure of privacy mental models. We believe mental models provide a framework for understanding user cognitive processing and reasoning and consequently privacy decison-making. In this paper we present an on-going study that use Amazon's Mechanical Turk and cognitive mapping technique to elicit and illustrate mental models. We compare the cognitive maps generated for two different questions and analyse their structural properties. We find that while a list of concrete privacy evaluations populate the cognitive maps when asked directly about privacy, the examples are generally scarce if not absent when queried about personal importance of the online environment. We also find that the degree of vertices complemented with the source and sink vertices can help to identify key concepts, triggering links and clusters within the maps.

privacy, mental models, cognition, user, decision-making, cognitive maps, usability

1. INTRODUCTION

Online privacy designs vary among categories of privacy-by-policy and privacy-by-architecture or privacy-by-design approach (Spiekermann and Cranor, 2009). Although legally and technically sound, these approaches are far from being effective as illustrated by the dichotomy between privacy attitudes and behaviour (Spiekermann et al., 2001; Acquisti and Grossklags, 2005).

We seek to explain this phenomenon by hypothesising that the design does not correspond to users' cognition in privacy desicion-making. In particular, perception, interpretation and evaluation of privacy decisions may exhibit cognitive associations not taken into account in the design for privacy. By providing an excerpt into human cognition, mental models provide us with a gateway to investigate cognitive associations. In addition, it enriches our ability to communicate with users in a manner that tunes into their mental models and activate privacy attitude associations. Furthermore examination of the content of users' privacy mental models, of the associations between concepts and their properties such as the proximity and similarity between clusters will provide key insights that support effective and usable design interventions.

This paper first provides the background research followed by presentation of a study conducted as part of our endeavour to develop users' mental models of online privacy. The study aims to answer a research question: *How do different framings of questions affect users' mental models of privacy?* We present the methodology followed by an analysis of the structure of the cognitive maps developed. We then discuss our findings and provide our future research directions.

2. BACKGROUND

In this section we introduce the difficulties in eliciting the cognitive dimensions of privacy. We provide a brief of mental models and their use in privacy reseach follows. The section ends with a note on the cognitive mapping technique.

2.1. Privacy Cognition

Previous research aimed towards understanding user privacy online has elicited privacy perceptions

and concerns, gathered self-reports of privacy behaviour or observed behaviour under laboratory design settings (Spiekermann et al., 2001; Acquisti and Grossklags, 2005). Elicitation of privacy perception and concern is a difficult task due to its sensitivity to priming effects. The methodology poses the risk of triggering cognitive associations, processing and activation of mental models leading to responses that might not usually arise in everyday interactions. This might explain the dichotomy observed by studies of privacy concerns and behaviour. Indirect approaches are thus often used such as eliciting perceived risks of online interactions (Miyazaki and Fernandez, 2001) or disclosure decision in specific context (Spiekermann et al., 2001).

2.2. Mental Models

Mental models are internalised, mental representations of a device or idea that facilitates reasoning (Johnson-Laird, 1983). They are simplistic and small-scale representations of reality (Craik, 1943). 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 (Larkin, 1983). Compared to novices, experts use chunking strategies to represent problems thus helping in problem representation (Chi et al., 1981; Chase and Simon, 1973).

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 (Norman, 1983). 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. However the *mental model uncertainty principle*, that is mental models are not directly accessible or observable, poses the inherent problem of representing mental models (Richardson et al., 1994).

2.2.1. Mental Models of Privacy

Mental models therefore promise a valuable framework to facilitate investigation. They do so by enabling illustration of conceptual relationships that hold semantic information, which would portray users' cognitive processing and reasoning. Mental models have been associated with privacy and security research before through analogies and metaphors. These include 'situational faces' (S.Lederer et al., 2003), 'audience-view' (Richter-Lipford et al., 2008), card-based metaphors (Wastlund et al., 2012), physical security model (Raja et al., 2011) or modeling of security risks (Camp, 2009). These involve areas of applications such as security warnings (Bravo-Lillo et al., 2011; Diesner et al., 2005) including firewalls (Raja et al., 2011), mobile security (Lin et al., 2012), end-to-end email security as means to e-mail protection (Renaud et al., 2014) and anonymous credentials (Wastlund et al., 2012). Recently there have also been proposals to elicit user mental models of security and privacy (Volkamer and Renaud, 2013; Coopamootoo and Groß, 2014).

2.3. Cognitive Mapping

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 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 (Tolman, 1948). However Axelrod (1976) used it as a 'map of cognition' while Eden later used it as reference to a map 'to aid cognition' (Eden, 1992). Axelrod's map of cognition has been used in artificial intelligence (Kosko, 1986) and experimental research such as system dynamics (Doyle and Ford, 1998). In our research, we also use cognitive maps as originally referred to by Axelrod. Also an agreed upon cognitive mapping methodology is not yet available between research domains (Vennix, 1990).

3. 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.

3.1. Design

Our main research question is: *What do user mental models of online privacy consist of?* Given that user reports of privacy are primed by the framing and design of studies, we postulate mental models

elicited are likely to suffer from such biases. The research question leads to subordinate questions such as: *How do different framings of questions affect users' mental models of privacy?*.

Our on-going research includes elicitation of privacy mental models. Given the presence of the researcher might affect responses provided, we opted to start the research with Amazon's Mechanical Turk for its accessible subjects. We are conducting between-subject studies with questions requiring 100 to 250 words responses. Our questions were:

- **Q1** - What does privacy online mean to you?
- **Q2** - What do you usually use the internet for? What is important to you when you are online?

The first two questions are chosen since they are believed to be at far ends of the direct to indirect spectrum. As a test case for our methodology we collected and analysed the data of five participants for each of these. We aim to add to the study with these questions framed differently such as positive and negative framings or questions including sharing rather than privacy for instance we recently launched another question: **Q3** - What does sharing online mean to you?

3.2. Elicitation Process

The free text collected by Mechanical Turk are first collated. We then use CMapTools to develop cognitive maps[1] made up of a hierarchy of concepts connected to each other via directional links as shown in Figure 1 and Figure 2. The elicitation process is as follows:

- each response is divided into *distinct phrases* of no more than 10–12 words long (possibly much shorter),
- statements within each phrase such as subject concept A exercising an 'action' on an object concept B are identified,
- relations are classified such as ontology ('is a', 'includes'), constraint (restriction of the application of the concept), cause-effect 'action' between the concepts or negation thereof,
- relative clauses associated with a concept already consumed are translated by duplicating the concept and forming a separate phrase,
- if concept A does not exist in teh concept set, a vertex labelled with concept A is created,

[1] `http://cmap.ihmc.us`

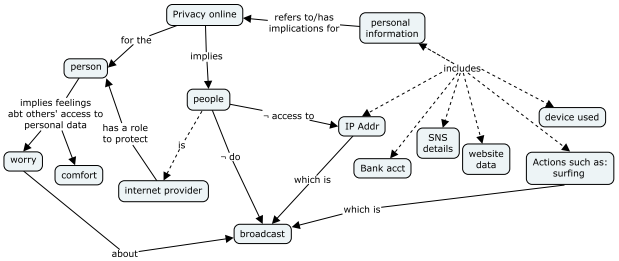

Figure 1: *Cognitive Map e.g. for* ***Q1****. (The dotted lines represent ontologies, the logical negation means 'not')*

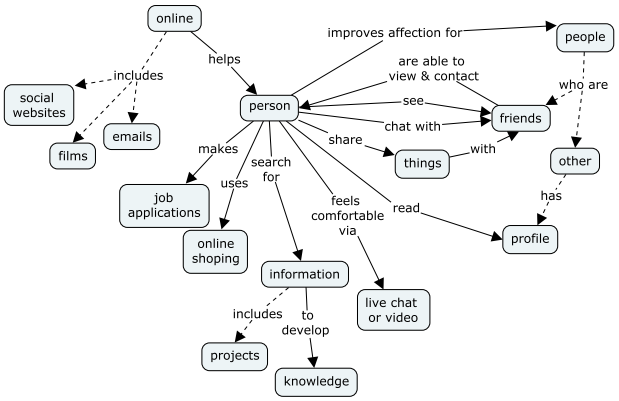

Figure 2: *Cognitive Map e.g. for* ***Q2****. (The dotted lines represent ontologies)*

- if concept B does not exist in teh concept set, a vertex labelled with concept B is created,
- an arrow from A to B labelled with the identified relation is created, typed with the relation type.

Definition 1 (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.*

3.3. Structural Analysis

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.

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	life	online life, job application
3	inviolable right	account, files, pictures, messages, friends
4	data, identity anonymity history	email, personal identifiers, posts, prying eyes
5	anonymous	person, people, criminals, email in an article or blog

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 source)
2	social websites, films, emails, job applic., online shopping, information, live chat or video	(no source)
3	better, perfect, high speed, travel tickets, bills, to know something, shopping, online work, films, favourite site	(no source)
4	hackers, banking sites, tv/movies, social media, email, information, quickly (gather information)	personal financial information
5	traffic load, power to connect, great (communicating with people), activities done online	social reach, topics

- for **Q2**, three maps had a shallower 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 shares specific information with and the benefits of obtaining information on the internet.

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**.

4. DISCUSSION

Sink and source together with degree of vertices to point important concepts or clusters for participants. These might be an indication of triggers to activate more elaborate mental models of privacy or enable privacy attitude evaluation. 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 however be interesting to find out whether the part of the mental model triggered links that approveor reject decisions and behaviour. 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'. 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 the participants' cognitive ability with respect to the question. However the shape can be influenced with thoughts and ideas that are more salient at the time of participating in the study or 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.

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' (Richardson et al., 1994) different elicitation approaches might lead to different

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)

results. In addition given the questionable stability of mental model over time, techniques to ascertain stability akin to those in trait theory would be valuable for the research. 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.

More extensive research are also needed to eliminate potential confounding explanations and investigate the multitude of framing possibilities. In addition the quality, readability and complexity of the questions and participants' cultural background are important confounds to the mental models derived.

5. CONCLUSION AND FUTURE WORK

In this paper we present initial results of on-going research aimed at depicting user mental models of online privacy. Our study was aimed at assessing cognitive maps produced from different framings of an elicitation question. We conclude that the way the question is designed influences the structural properties of the mental models gathered. In addition we find that the methodology presents the potential to contribute towards identifying users' privacy inclinations and their cognitive ability with respect to privacy. For instance we posit that the domain score complemented with the head and tail nodes can help to identify key concepts leading to most associations thus potentially helping to categorise privacy concerns.

Our future work first includes expansion of our study across a larger user pool while investigating an array of different questions that could have privacy implications and hence generate privacy related mental models. We aim to facilitate this by developing a structured method of eliciting, analysing and mapping users' cognitive maps. We consider comparing results of cognitive mapping with other approaches such as repertory grids or semiotic analysis.

Second, the investigation will benefit from other methods of collecting user data such as face-to-face interviews, drawing of concept maps or observations. Analysis methods such as co-occurence matrix, cognitive distance, cluster analysis, multidimensional scaling and hierarchical cluster analysis might shed more light on the cognitive maps depicted.

Third we aim to explore a series of subordinate research questions and hypotheses such as:

- whether and how mental models support behaviour,
- how cognitive ability influence privacy mental models,
- how cognitive effort impact the activation of models and their size, shape, complexity,
- how the type of reasoning such as inductive or deductive methods influence privacy models.

REFERENCES

Acquisti, A. and J. Grossklags (2005). Privacy and rationality in individual decision making. *IEEE Security & Privacy 2*, 24–30.

Axelrod, R. (1976). *The structure of decision*. Princeton University Press, N.J.

Bravo-Lillo, C., L. Cranor, J. Downs, and S. Komanduri (2011, March). Bridging the gap in computer security warnings: A mental model approach. *Security Privacy, IEEE 9*(2), 18–26.

Camp, L. (2009). Mental models of privacy and security. In *IEEE Technology and society magazine*, Volume 28.

Chase, W. G. and H. A. Simon (1973). Perception in chess. *Cognitive psychology 4*(1), 55–81.

Chi, M. T., P. J. Feltovich, and R. Glaser (1981). Categorization and representation of physics problems by experts and novices*. *Cognitive science 5*(2), 121–152.

Coopamootoo, K. P. L. and T. Groß (2014). Mental models for usable privacy: A position paper. In T.Tryfonas and I.Askoxylakis (Eds.), *HAS 2014*, Volume 8533 of *LNCS*, pp. 410–421. Springer Int.

Craik, K. (1943). *The nature of explanation*. Cambridge University Press.

Diesner, J., P. Kumaraguru, and K. M. Carley (2005). Mental models of data privacy and security extracted from interviews with indians. In *55th Annual Conference of the International Communication Association (ICA), New York, NY*.

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

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

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

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

Larkin, J. (1983). *Mental models*, Chapter Expert representations of physics problems. Lawrence Erlbaum Associatesn.

Lin, J., S. Amini, J. I. Hong, N. Sadeh, J. Lindqvist, and J. Zhang (2012). Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pp. 501–510. ACM.

Miyazaki, A. and A. Fernandez (2001). Consumer perceptions of privacy and security risks for online shopping. *Journal of Consumer Affairs 35*, 27.

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

Raja, F., K. Hawkey, S. Hsu, K.-L. Wang, and K. Beznosov (2011). Promoting a physical security mental model for personal firewall warnings. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pp. 1585–1590. ACM.

Renaud, K., M. Volkamer, and A. Renkema-Padmos (2014). Why doesn't jane protect her privacy? In *Privacy Enhancing Technologies*, pp. 244–262. Springer.

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

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

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

Spiekermann, S. and L. Cranor (2009). Engineering privacy. In *IEEE Trans on S/w Eng*, Volume 38, pp. 67–82.

Spiekermann, S., J. Grossklags, and B. Berendt (2001). E-privacy in 2nd generation e-commerce: privacy preferences versus actual behavior. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pp. 38–47. ACM.

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

Vennix, J. (1990). Mental models and computer models: Design and evaluation of a computer-based learning environment for policy-making.

Volkamer, M. and K. Renaud (2013). Mental models–general introduction and review of their application to human-centred security. *Number Theory and Cryptography*, 255–280.

Wastlund, E., J. Angulo, and S. Fischer-Hubner (2012). Evoking comprehensive mental models of anonymous credentials. In LNCS (Ed.), *Open problems in network security*, Volume 7039, pp. 1–14.