BOISE STATE UNIVERSITY

Understanding Children's Mental Models of Privacy based on the Theory of Contextual Integrity

Dr. Hoda Mehrpouyan
Co-authors:
Jerry Alan Fails, Dhanush Kumar Ratakonda
hodamehrpouyan@boisestate.edu
Computer Science Department
2nd Symposium on Applications of Contextual Integrity

Proposed Approach

- We extend the concept of *contextual integrity* to provide mathematical models and algorithms that enables the creations and management of privacy norms for individual users.
  - The extension includes the augmentation of *environmental variables*, i.e. time. date, etc. as part of the privacy norms, while introducing an *abstraction and a partial relation over information attributes*.

The Proposed Framework

- The proposed framework is based on two sets of formal models:
  1. *User's Information Sharing Model (UISM):* represents the information sharing activities in real-time,
  2. *Privacy-Preserving Model (PPM):* formally specifies the user's privacy requirements.

Privacy verification is performed by mapping each action in UISM to its corresponding action in the PPM. In the case of not being able to map an action a privacy violation is detected and reported to user to get confirmation.

User Information Sharing Model (UISM)

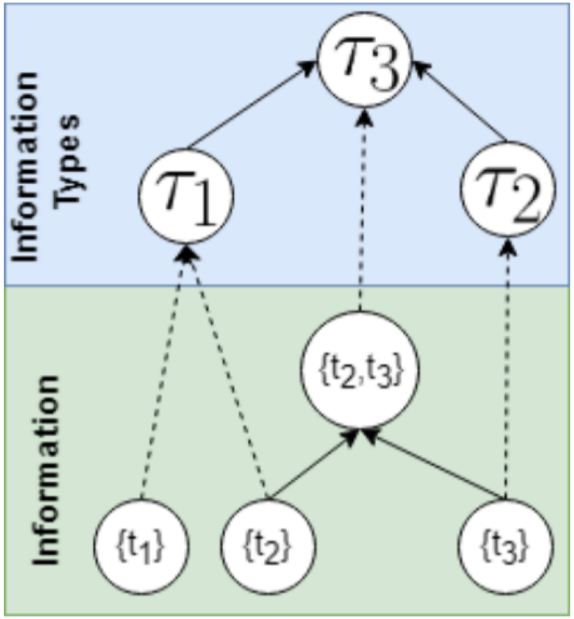
- UISM is designed based on the formal definition of:
  - *entities* that construct Information Communication mechanism based on *agent*
    - Model user's information sharing behavior with the recipients, which are defined as agents.
      - P is defined as a set of agents that are the recipient of the information sent from the user.
  - T is a set of *attributes* that defines the information shared with p ∈ P such as "home address" or "credit card number".
  - a *knowledge* state κ is defined as a set of tuples of the form (p , $\{t_1 \ldots t_k\}$) which describes the attributes $t_i \in T$ that is shared with an agent p

Abstractions

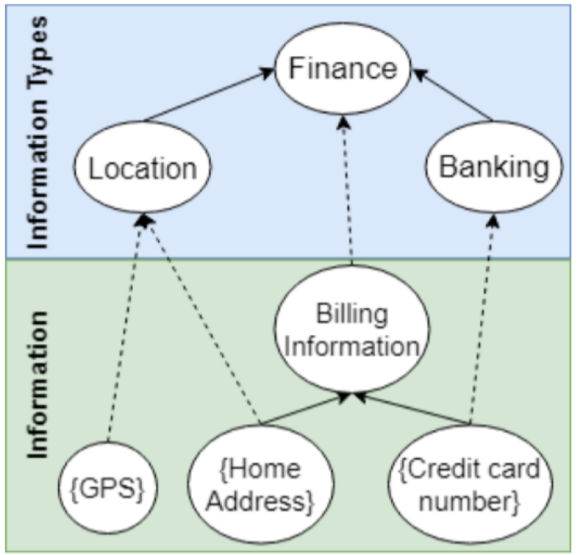
- In a user-centric approach is inefficient to define a separate privacy norm for each ρ (*role*) and τ (*attribute type*), the proposed model abstracts these two elements
  - This abstraction allows to have the same information disclosure norms with a set of agents or disclose a collection of attributes in a similar manner.
    - For example, the user could share her current location with all transportation applications, or the user could share her credit and debit cards' numbers with her close family members.

Attribute Type

Let τ be a set of *attribute types* and let AT be a partial map



(a)



(b)

Role Abstraction

- A set of *roles* (R) is defined to be assigned to an *agent* (p).
  - An agent can be assigned to multiple roles and roles are partially ordered based on their implication relation of their semantics.
  - The partial order $\leq on\ R$ is predefined as an input to the model, such that the role, ρ_1, “close friend" implies the role, ρ_2, "friend”$\rho_2 \leq \rho_1$.
    - The order between roles implies the amount of relative privacy restriction of them where $\rho_2 \leq \rho_1$ means that ρ_2 is more restrictive compared to ρ_1

Formal Definition of UISM

DEFINITION 1. ***(The User Information Sharing Model (UISM)***

Let $UISMM = (K, Act, \rightarrow, \kappa_0)$ *be a 4-tuple transition system where:*

- K *is a finite set of knowledge states* κ.
- $\kappa_0 \in K$ *is the initial state* $\kappa_0 = \emptyset$ *(no initial disclosures).*
- Act *is a set of communication actions.*
- $\rightarrow \subseteq K \times Act \times K$ *is a transition relation, transform the system state with actions* $(a, p, \widetilde{t})$ *as follows:*
  - $\kappa \xrightarrow{(sh, p, \widetilde{t})} \kappa'$, *where* $\kappa' = \kappa \cup \{(p, \widetilde{t})\}$,
  - $\kappa \xrightarrow{(st, p, \widetilde{t})} \kappa'$, *where* $\kappa' = \kappa \setminus \{(p, \widetilde{t'}) \mid \widetilde{t} \cap \widetilde{t'} \neq \emptyset\}$.

Privacy-Preserving Model (PPM)

- The Privacy-Preserving Model is designed to manage and govern user's information sharing activities at run-time.
  - based on the proposed UISM in the previous section, PPM model is required to govern the transitions between knowledge states according to the *norms* that the user specifies

Access Permissions

- $\mathcal{A}$ *is a subset of* $\mathcal{R}\times\tau$ such that if $(\rho,\tau)\epsilon\mathcal{A}$ then all agents with role ρ are allowed to access attributes with type τ
- What about *environmental conditions and temporal conditions ?*
  - We introduce the logic for *environmental conditions* ψ and *temporal conditions* φ to the definition of the privacy norm.
    - In this model, environmental conditions are represented a set of variables V, where each v$\in V$ describes the state of an environment such as system's time, day and other attributes.

$$pred_I ::= v \leq n \mid v < n \mid v == n, v \in V_I, n \in \mathbb{Z}$$

$$pred_B ::= v \mid true \mid false,\ v \in V_B$$

$$\psi ::= \neg\psi \mid \psi \wedge \psi \mid \psi \vee \psi \mid pred_i, \forall V_i \in V$$

Access Permissions II

- Temporal Conditions φ
  - define the precedence of two communication actions or a constant occurrence a communication actions can be sufficiently defined by the concatenation and Kleen star operations over $\mathcal{A}$ (the alphabet):

$$\varphi, \phi ::= (\rho, \tau) \mid \varphi \cdot \phi \mid \varphi^*, (\rho, \tau) \in \mathcal{A}$$

$$\varphi = \mathcal{A}_1^* \cdot ((\rho, \tau_1) \cdot \mathcal{A}_1^* \cdot (\rho, \tau_2))^* \cdot \mathcal{A}_1^*$$

  - the repetition of an event up to a constant k times could be expressed with the following formula, where the power operator describes the number of times a regular expression should be repeated

$$\varphi = \mathcal{A}_2^*((\rho, \tau) \cdot \mathcal{A}_2^*)^k$$

Norms

- In this research, *norms* are the formal definition of user's privacy requirements that are used to govern user's information sharing behavior.
  - In order to minimize the risk of unwanted information sharing, we assume that if an action is not explicitly defined as part of the user's privacy policies then it is ***forbidden***.
  - Therefore, the only type of norms that the user defines are ***positive norms***, i.e., allowed norms.

$$((\rho, \tau), \psi, \varphi,)$$

Consistent Norms

- Two norms are consistent when $n_1 = ((\rho_1, \tau_1), \psi_1, \varphi_1)$and $n_2 = ((\rho_2, \tau_2), \psi_2, \varphi_2)$ are consistent when one of the four consistency conditions holds:

C1. $\nexists p \in \mathcal{P} : \rho_1 \in AR(p) \wedge \rho_2 \in AR(p)$, *that is, the norms defined for the roles with no common agents.*

C2. $\nexists \tilde{t} \in \mathcal{P}(T) : AT(\tilde{t}) \leq \tau_1 \wedge AT(\tilde{t}) \leq \tau_2$, *that is, norms are defined for attribute types with no common information attribute*

Consistent Norms II

- Same role should be used in the access permission and the sequencing condition of a norm.
- An attribute type and its children are not allowed to exist in the same regular expression

C3 . $\rho_1 < \rho_2$ *and either* $\tau_1 \leq \tau_2$ *or* $\tau_2 \leq \tau_1$ *then* $\psi_1 \implies \psi_2 \wedge \mathcal{L}_{\downarrow}(\varphi_1)_{\varphi_2} \subseteq \mathcal{L}_{\downarrow}(\varphi_2)_{\varphi_1}$, *that is,* n_2 *is for a specialized role* ρ_2 *of* ρ_1 *and its attribute type* τ_2 *encompasses* τ_1 *or vise verse then environmental condition of* ψ_2 *should be the same or less restrictive than of* ψ_1 *and its regular expression* φ_2 *should describe the same or less restricted projected language than of* φ_1

C4 . $\rho_1 < p > \rho_2$ *or* $\tau_1 < t > \tau_2$ *then* $\psi_1 \Leftrightarrow \psi_2 \wedge \mathcal{L}_{\downarrow}(\varphi_1)_{\varphi_2} = \mathcal{L}_{\downarrow}(\varphi_2)_{\varphi_1}$. *If there is at least one agent that can be assigned to both unrelated roles or an information attribute that share a common child then the environmental conditions and the projected language of the regular expressions must be equivalent.*

Policy Compliance Verification

DEFINITION 4. ***(Privacy-Preserving Model)*** *is a set of observers over norms $\mathcal{N}$ where each observer is a tuple of $(\widehat{K}, \widehat{Act}, c, m)$ representing $n_i = ((\rho, \tau), \psi, \varphi) \in \mathcal{N}$ where $\widehat{K} = (\rho, \tau)$, $c = \psi$ is the pre-condition and m is a monitor representing φ regular expression. The transition $\widehat{Act}$ is given to Monitor m to update the state of the monitor.*

To ensure that the user's behavior is compliant with the privacy policy, we need to map the current state and the next state of user's behavior model to the privacy preserving behavior model.

DEFINITION 5. ***(Mapping from user behavior to privacy preserving domain)*** *Let $MS : K \rightarrow \widehat{K}$ be a surjective function, where $MS(p, \widehat{t}) = \{(\rho, \tau) | \rho = AR(p), \tau = AT(\widetilde{t})\}$ and $MT : Act \rightarrow \widehat{Act}$ where:*

$$MT(a, p, t) = \{(\rho, \tau) | \rho \in AR(p) \wedge \tau \in AT(t)\} \; if \; a = sh$$

Children's Understanding of Privacy

Semi-structured Interviews with children, age 7 to 11:

- In our study, we assumed three contexts of family (parents, siblings and family friends roles), school (teacher, classmates roles), friends (close friends, friends, online friends roles), and neighbors

Child interview structure

Segment 1. Define thirteen privacy and security-related words to the best of their knowledge (See Figure 1). The words were chosen based on the common terminologies that are mostly used in the privacy and security aware tutorials and interactive books such as Cyberheroes [8]. The focus of the study was more on privacy threats related to data tracking.

Segment 2. What data they are comfortable sharing data (e.g., first name, full name, picture, etc.) with different roles (e.g., neighbors, family friends, online friends). The complete matrix (see Figure 2), few cells were grayed out for obvious reasons. This evaluates childrens' understanding of roles and information attributes in CI.

Segment 3. Additional questions to understand children's comfort in sharing information in a verbal conversation or through email/electronic media. We asked them to define privacy threats, security threats, and to share experiences with them. This evaluates childrens' understanding of transmission principles in CI.

	Can Describe?	
Privacy Concept	Yes	No
Anonymity	0	6
Cyber crime	2	4
Cyber Disguise	2	4
Cyber Invisibility	2	4
Cyber trail	2	4
Data tracking	1	5
Digital Footprint	1	5
Encryption	0	6
Hacking	0	6
Pop-up	1	5
Scam	3	3
Secret	5	1
Secure connection	3	3

Figure 1: Privacy concepts that children were asked to define (to the best of their ability).

Result

	Parents	Siblings	Teachers	Close friend	Classmates	Friend	FamilyFriends	Neighbors	Online friend
First name								4	
Full name				5	5	4	6	3	1
Your picture			2	5	4	5	5	3	1
Email address		4	3	2	1	1	2	1	0
Home address			2	4	5	2	5		0
Age				5	5	6	6	3	2
Birthday				5	5	6	4	3	1
Birthplace			3	4	4	2	2	1	1
Pet's name			1	3	3	3	3	3	3
Telephone #			1	4	3	1	4	3	1
Nick Name			2	5	4	5	5	2	2
Username	5	4	2	2	1	1	1	1	1
Password	4	3	2	0	1	0	1	0	0

Legend	0	1	2	3	4	5	6	NA

Conclusion and Future Works

- Current privacy management systems cannot address dynamic requirements of privacy adequately since they are not designed based on the users' privacy perspectives
- To overcome these limitations, the contextual integrity theory has been customized to address the privacy needs of individual users.
- The future work will eliminate the current user interface and user's privacy norms will be generated automatically utilizing text analysis, speech recognition, and AI algorithms that can infer user's privacy policies based on the user's relationships and information sharing behaviors

Our Papers:

- Rezvan Joshaghani, Stacy Black, Elena Sherman, and Hoda Mehrpouyan. 2019. Formal specification and verification of user-centric privacy policies for ubiquitous systems. In Proceedings of the 23rd International Database Applications & Engineering Symposium (IDEAS '19). ACM, New York, NY, USA, Article 31, 10 pages. DOI: https://doi.org/10.1145/3331076.3331105
- Stacy Black, Rezvan Joshaghani, Dhanush kumar Ratakonda, Hoda Mehrpouyan, and Jerry Alan Fails. 2019. Anon what what?: Children's Understanding of the Language of Privacy. In Proceedings of the 18th ACM International Conference on Interaction Design and Children (IDC '19). ACM, New York, NY, USA, 439-445. DOI: https://doi.org/10.1145/3311927.3325324

Thank you,

Dr. Hoda Mehrpouyan

hodamehrpouyan@boisestate.edu

Q&A?