

PERC_LAB Home of Privacy Engineering – Regulatory Compliance Research



Lattice-based Contextual Integrity Analysis of Social Network Privacy Policies

Stephen Kaplan, Dylan Bulmer, Avery Gosselin, & Sepideh Ghanavati 1 October 2021

Lattice-based Contextual Integrity Analysis

- Lattice-based Contextual Integrity Analysis (LCIA) is a four-phase privacy policy analysis framework which aims to:
 - Identify and quantify ambiguity within online social network (OSN) privacy policies
 - Evaluate and rank the privacy practices of OSNs
 - Allow us to make predictions about how likely an OSN's privacy policy is to mislead users with regard to its information flow practices relative to other OSNs



Problem Description

Privacy Policy	ge				
When you use our services, you're trusting us with your information. We understand this is a big					
responsibility and work hard to protect your information and put you in control.	oses	cample,			
This Polyany Policy is money to help you understand what information we collect who we called it and have	for	site			
This Privacy Poincy is mean to help you understand what information we conect, why we conect it, and now	an	ing			
you can update, manage, export, and delete your information.					
Privacy Checkup					
Looking to change your privacy settings?		sexual	Im	of	
	hem				
Take the Privacy Checkup	ntify	ne or			
We build a range of services that help millions of people daily to explore and interact with the world in new		ct the			
ways. Our services include:	y	er			0
			or		
Apps, sites, and devices, like Search					
Products that are integrated into third-party apps and sites, like ads and embedded Maps					
You can use our services in a variety of ways to manage your privacy. For example, you can sign up for					
a Account if you want to create and manage content like emails and photos, or see more relevant search	ne	we			
results. And you can use many services when you're signed out or without creating an account at all, like		e data			
searching on or watching YouTube videos. You can also choose to browse the web privately using Chrome	our	ins. We			С
in Incognito mode. And across our services, you can adjust your privacy settings to control what we collect		alytics,	1		
and how your information is used.		om that			
To help explain things as clearly as possible, we've added examples, explanatory videos, and definitions for	. For				
key terms. And if you have any questions about this Privacy Policy, you can contact us.	red		s		
			0.05	in	
Information we collect			5112,		
We want you to understand the types of information we collect as you use our services	ogle	e may	the state		
We collect information to provide better services to all our users - from figuring out basic stuff like which		ount	iy	hd	
Brivaov controle					
Privacy controls		,sues	ple,	s	18
you might be facing					
· · · · · · · · · · · · · · · · · · ·				t.	
your ad settings.					
such as browser web storage or application data caches, databases	, and server log	5.			
My Activity					

PERC_LAB Home of Privacy Engineering -Regulatory Compliance Research

- Privacy policies are often long and confusing
- Difficult for users to know exactly what information an OSN collects
- More difficult for users to know exactly how their information is used and shared

Background I

- LCIA relies on the Contextual Integrity framework
 - Nissenbaum, H. "Privacy in Context: Technology, Policy, and the Integrity of Social Life." Stanford University Press (2009)
 - Shvartzshnaider, Y., Apthorpe, N., Feamster, N., & Nissenbaum, H. "Going against the (Appropriate) Flow: A Contextual Integrity Approach to Privacy Policy Analysis." AAAI (2019)

[Attribute] [Subject & Sender] We encrypt all of the information that we collect from you. [Transmission [Recipient] Principle]



Background II

- LCIA relies on lattice representations of privacy practices
 - Ghazinour, K., Majedi, M., and Barker, K. "A Lattice-Based Privacy Aware Access Control Model." 2009 International Conference on Computational Science and Engineering (2009)



Research Questions

- *RQ-1.* What are the similarities and differences in the way existing OSNs define, protect, and violate user privacy?
- *RQ-2.* How can OSN privacy practices be compared in a standardized way?

RQ-3. What relationships exist between poor OSN privacy practices, poor privacy policies, and gaps in user understanding of privacy?



LCIA Methodology



Phase 1: Privacy Practice Identification





Phase 2: Contextual Integrity Analysis





Phase 3: Lattice Analysis





Phase 4: Rank and Trend Analysis

- Sum the normalized CI_{Score} and L_{Score} of each OSN to form a combined LCI_{Score}
- Rank the OSNs according to their LCI_{Score}
- Identify trends in the analyzed sample of OSNs



Preliminary Analysis - Data Collection

Category	OSNs	
General	Facebook, Twitter, Reddit, Tagged, VK	
Health	Samsung Health, CaringBridge	
Image Sharing	Instagram, Imgur, Flickr, Pinterest, DeviantArt, *Ello.co, *PixelFed, We ♥ It	
Video Sharing	YouTube, Twitch, Vimeo, TikTok	
Dating	Tinder, Grindr, Match, Bumble	
Blogging	Tumblr, Blogger, Quora, OpenDiary	
Music Sharing	SoundCloud, MySpace	
Text Sharing	Goodreads, Wattpad	
Professional Networking	LinkedIn, NearPeer	
Voice Chat	Discord, Skype, Microsoft Teams, TeamSpeak	
Messaging	WhatsApp, Facebook Messenger, Snapchat, Slack, Moco, *Mastodon, *Element, *Signal, *Telegram	
Content Discovery	Mix	
Business Discovery	Yelp, FourSquare	
Gaming	Habbo	

- Compiled a list of 50 social networks in 14 categories
- Filtered for networks supporting n-removed connections
- Randomly selected n samples from each category
- Obtained each OSN's privacy using using our *PolicyAccumulator*

Preliminary Results

OSN	Phase 2	Phase 3	Phase 4
VK	0.97	0.30	0.78
Facebook	-1.88	-0.99	-1.76
CaringBridge	0.76	1.16	1.18
DeviantArt	-1.22	1.59	0.23
PixelFed	0.71	0.73	0.88
YouTube	-1.78	-1.85	-2.23
Tumblr	-0.31	0.30	-0.01
SoundCloud	0.10	-0.99	-0.55
Wattpad	0.61	0.30	0.56
LinkedIn	0.46	0.30	0.47
Snapchat	0.15	0.73	0.54
Yelp	0.86	-0.99	-0.08
Habbo	0.56	-0.56	0

YouTube (Least Privacy-Preserving) Facebook SoundCloud Yelp Tumblr Habbo DeviantArt LinkedIn Snapchat Wattpad VK PixelFed CaringBridge (Most Privacy-Preserving)

Conclusion

- We presented a four-phase privacy policy analysis framework
 - Systematically compares the privacy practices of OSNs

- We demonstrated LCIA's potential effectiveness
 - Performed a preliminary evaluation of LCIA on 13 OSN's privacy policies
 - Ranked social networks based on overall privacy practices, revealing cases where users may misunderstand privacy practices



Future Work

- Leverage unsupervised machine learning in annotation process
- Conduct a user study on users' perception of privacy violation
- Evaluation of larger dataset to reveal generalizable insights
- Implementation
 - Real time analysis of privacy policies
 - Policy analysis prior to application publication



Thank you!

Slides will be available at <u>skaplan.io/LCIA</u> and <u>dylanbulmer.com/publications/LCIA</u>

Reach out with any questions

stephen.kaplan@maine.edu



PERC_LAB

Home of Privacy Engineering – Regulatory Compliance Research

