The robustness of security and privacy properties in decentralized applications

Guillaume Piolle

The robustness of security and privacy properties in decentralized applications

Guillaume Piolle
guillaume.piolle@centralesupelec.fr
https://guillaume.piolle.fr/

CentraleSupélec / Inria – CIDRE research group (Rennes)

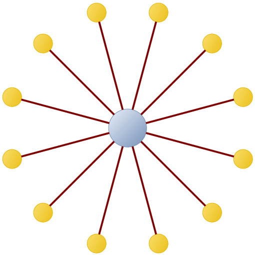
Surveillance, Resilience and Privacy
december 2018

The CIDRE research group

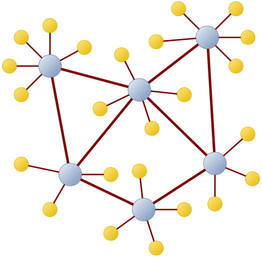
Common research group between CentraleSupélec, Inria, CNRS and Rennes 1 university.
Domains of interest:

- **Information security**: threat analysis (malware analysis, physical attacks, attacks on blockchains, security visualization. . .), intrusion detection (most notably via information flow monitoring, on Linux, Android and distributed systems) and proactive security (formally specified systems, reactive frameworks for virtualization, specialized blockchain models. . .) ;
- **Privacy and personal data protection**: policy-driven / privacy-preserving architectures and protocols, computational law, geo-privacy, privacy-preserving ML. . .

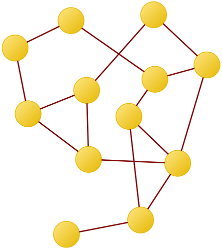
A taxonomy of decentralization



Centralized architecture



Decentralized architecture

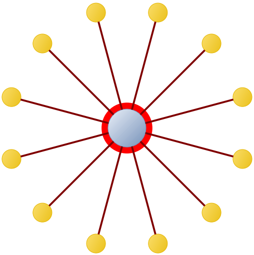


Distributed architecture

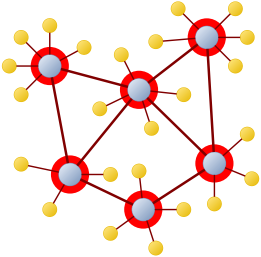
Commonly accepted hypothesis

Decentralization is a good move for privacy, it increases user control and avoids “Big Brothers”.

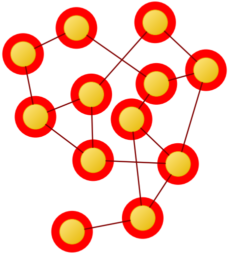
What does an adversary need to compromise?



Only one node



A set of nodes



All the nodes

We consider *blanket attacks* rather than *targeted attacks*

Informally, a higher degree of decentralization seems to make the system more *robust* to adversaries of a given capability level.

An example of partial decentralization

Starting point: a trustworthy service, easy to use. . .

A provider is running a centralized encrypted file hosting service. Users push their files to the server. They manage encryption and decryption through a web interface, the provider serving as a key escrow *(ew. . .)*.

An example of partial decentralization

Starting point: a trustworthy service, easy to use...

A provider is running a centralized encrypted file hosting service. Users push their files to the server. They manage encryption and decryption through a web interface, the provider serving as a key escrow *(ew...)*.

A step towards decentralization

In order to "improve privacy and user empowerment", the company updates their service to allow encrypted data to be stored on the users' devices, rather than on their cluster.

An example of partial decentralization

Starting point: a trustworthy service, easy to use...

A provider is running a centralized encrypted file hosting service. Users push their files to the server. They manage encryption and decryption through a web interface, the provider serving as a key escrow *(ew...)*.

A step towards decentralization

In order to "improve privacy and user empowerment", the company updates their service to allow encrypted data to be stored on the users' devices, rather than on their cluster.

Application decentralization does not seem to be monolithic

Data storage is potentially **distributed**, but key management, and therefore encryption/decryption capabilities, remain **fully centralized**...

Per-property decentralization levels

It is possible (and useful) to evaluate the level of decentralization of a system, **feature by feature** and **property by property**, allowing for a better grasp of its security and privacy properties.

This can open the way to metrics for the quantization of a system's architectural **robustness to faults and attacks**.

To get this, we need:

- A **scale** for the level of decentralization (for starters, the one we have will do the trick);
- A **definition** of what it means, for each property or feature, to be decentralized at a given level.

An instance of a privacy property: unlinkability

Unlinkability as defined in the ISO/IEC *Common Criteria*

[Unlinkability] ensures that a user may make multiple uses of resources or services without others being able to link these uses together.

Unlinkability *implies* (and therefore is stronger than) anonymity and pseudonymity: in theory, **you can be fully anonymous in a system, and yet remain traceable**.

It is a prerequisite to other privacy-damaging capabilities:

- The ability to **single-out** a user's information in a mass of data, to **attribute** it;
- The ability to **track** a user in an information stream;
- The ability to perform various kinds of **surveillance**, for instance by systematically clustering / attributing / discriminating data and profiles;
- . . .

Quantifying the degree of decentralization of unlinkability

- **Centralized unlinkability:** the users' actions are unlinkable for everyone but a central authority.
  Only the central authority needs to be compromised in order to link the actions of all users;
- **Decentralized unlinkability:** the users' actions are unlinkable for everyone but one or several entities among a given set of autonomous authorities.
  The set of all autonomous authorities needs to be compromised in order to link the actions of all users;
- **Distributed unlinkability:** the users' actions are unlinkable for everyone but the actor.
  The set of all peers needs to be compromised in order to link the actions of all users.

R. P. M. Marin, *Enhancing Privacy Protection in Social Network Systems Through Decentralization and Policy Conflict Management*, PhD thesis, CentraleSupélec, 2015.

Application to social networking systems

	Facebook					SuperNova					Diaspora					PrivacyWatch					PeerSoN					Safebook					FOAF				
	?	*0*	*C*	*D*	*FD*	*?*	*0*	*C*	*D*	*FD*	*?*	*0*	*C*	*D*	*FD*	*?*	*0*	*C*	*D*	*FD*	*?*	*0*	*C*	*D*	*FD*	*?*	*0*	*C*	*D*	*FD*	*?*	*0*	*C*	*D*	*FD*
Privacy-related Properties																																			
Architectural Services																																			
Retrieval			C							FD				D						FD					FD					FD					FD
Communication			C							FD				D						FD					FD					FD					FD
Search			C						D					D				C						D						FD					FD
Storage																																			
Storage Space			C							FD				D						FD					FD					FD			C		
Replication			C						D					D						FD					FD					FD	?				
Data Suppression		0				?								D						FD	?					?					?				
Privacy Policy Management																																			
System Policy Administration			C				0							D						FD		0					0					0			
Peer Policy Administration			C							FD				D						FD					FD					FD			C		
System Policy Enforcement			C				0							D						FD		0					0					0			
Peer Policy Enforcement			C							FD				D						FD					FD					FD			C		
Security Aspects of Privacy																																			
Data encryption		0								FD		0								FD					FD			C			?				
Traffic encryption			C			?								D		?					?							C					C		
Anonymity			C						D					D					D					D				C				0			
Pseudonymity			C						D					D					D					D				C				0			
Unlinkability			C						D					D				C						D				C					C		
Unobservability			C						D					D				C						D						FD			C		

An attempt at characterizing resilience

Can we use the concept of **privacy property decentralization** to contribute to a definition of **resilience to privacy breaches**?

The multi-headed concept of (cyber-) resilience

One possible definition: US Presidential Policy Directive PPD-21 (Critical Infrastructure Security and Resilience)

> *The term "resilience" means the ability to **prepare for** and **adapt to** changing conditions and **withstand** and **recover** rapidly from disruptions.*

An attempt at defining resilience

Aspects with relevance to our metrics

- **Preparation:** making architectural design choices based on such evaluations may pertain to Privacy/Security by Design (or proactive security);
- **Withstanding:** also referred to as business continuity in other frameworks. Going towards distributed security/privacy properties makes it less likely for a breach to impact a lot of users, thus favouring (degraded) service availability.

Ok, so the service remains available, but for whom? It's only one of the aspects not captured by those metrics...