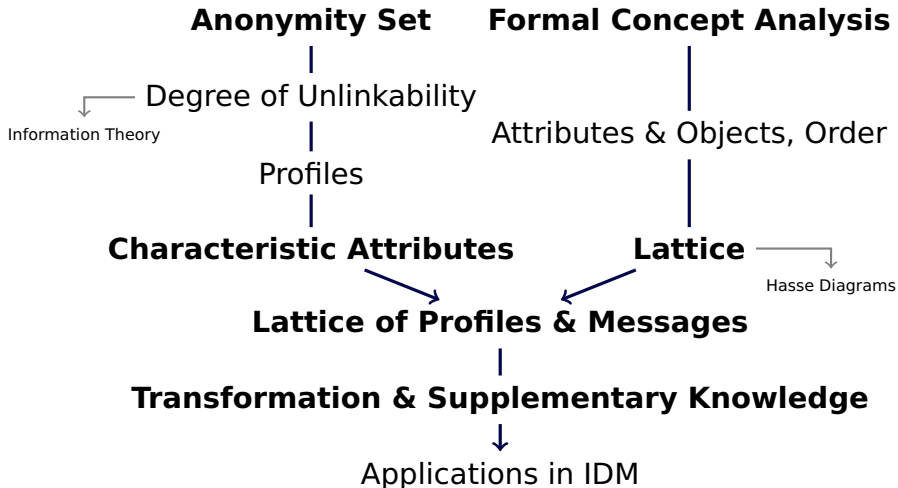


Linkability Estimation Between Subjects and Message Contents Using Formal Concepts

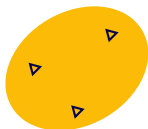
Stefan Berthold

Department of Computer Science
Technische Universität Dresden
01062 Dresden, Germany

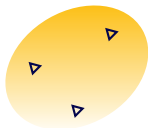
November 2, 2007



Anonymity Set Notion



anonymity set

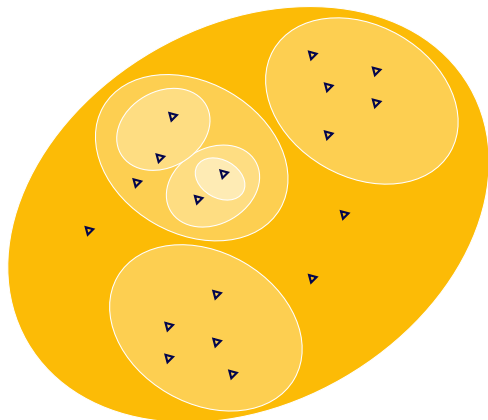


degree of anonymity

Historical development:

- Anonymity sets: size as measure of anonymity
- Unlinkability: items are more or less related
- Information theory: entropy as measure of anonymity

Profiles & Characteristic Attributes



Profiles

- Characteristic set of attributes
- Subjects can be assigned
- Profile refinements

Formal Concept Analysis

		carnivore	mammal	domesticated	herbivore	nonvolant
		<i>c</i>	<i>m</i>	<i>d</i>	<i>h</i>	<i>n</i>
Eagle	<i>E</i>	x				x
Lion	<i>L</i>	x	x			
Cat	<i>C</i>	x	x	x		
Horse	<i>H</i>		x	x	x	
Batman	<i>B</i>		x			x

$$\mathfrak{B} = \{(\{E, L, C, H, B\}, \emptyset),$$

$$(\{L, C, H, B\}, \{m\}),$$

$$(\{E, B\}, \{n\}),$$

$$\dots$$

$$(\{H\}, \{m, d, h\}),$$

$$(\{L, C\}, \{c, m\}),$$

$$(\emptyset, \{c, m, d, h, n\})\}$$

$$(\{L, C\}, \{c, m\}) \leq (\{L, C, H, B\}, \{m\})$$

$$\{L, C\} \subseteq \{L, C, H, B\}$$

$$\{c, m\} \supseteq \{m\}$$

Definition (Context)

Binary relation between objects and attributes.

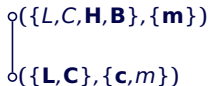
Definition (Concept)

Tuple (A, B) , that is a set of objects A and a set of attributes B , such that $A'' = B'$.

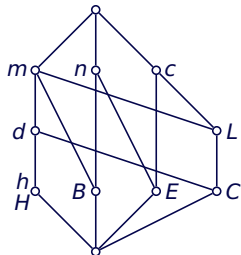
Definition (Concept Lattice)

Set of concepts \mathfrak{B} and lattice order \leq .

Graph-based Representation



subconcept–superconcept relation



entire concept lattice

Objects

<i>E</i>	Eagle
<i>L</i>	Lion
<i>C</i>	Cat
<i>H</i>	Horse
<i>B</i>	Batman

Attributes

<i>c</i>	carnivore
<i>m</i>	mammal
<i>d</i>	domesticated
<i>h</i>	herbivore
<i>n</i>	nonvolant

$$(\{L, C\}, \{c, m\}) \leq (\{L, C, H, B\}, \{m\})$$

Definition (Line diagram)

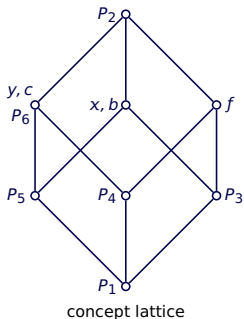
Hasse diagram with respect to the object sets.

Definition (Reduced LD)

Line diagrams, but omit

- redundancy in labels and
- arrow tips.

Lattice of Profiles



	seen at X	seen at Y	trade books	trade cars	trade food
Pseudonym 1	x	y	b	c	f
Pseudonym 2	x	x	x	x	x
Pseudonym 3	x		x		x
Pseudonym 4		x		x	x
Pseudonym 5	x	x	x	x	
Pseudonym 6		x		x	

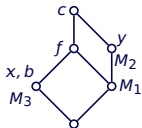
Definition (Objects)

Profiles or pseudonyms.

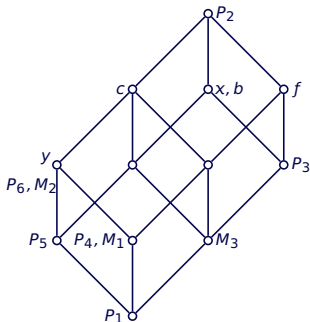
Definition (Attributes)

Data items and behavior.

Lattice of Messages



concept lattice of message sub-context



concept lattice of merged context

	seen at X	seen at Y	trade books	trade cars	trade food	
Message 1	M_1	x	y	b	c	f
Message 2	M_2		x		x	
Message 3	M_3	x		x	x	x
Pseudonym 1	P_1	x	x	x	x	x
Pseudonym 2	P_2					
Pseudonym 3	P_3	x		x		x
Pseudonym 4	P_4		x		x	x
Pseudonym 5	P_5	x	x	x	x	
Pseudonym 6	P_6		x		x	

Definition (Objects)

Messages or transaction pseudonyms.

Definition (Attributes)

Data items and behavior.

Incorporation of Supplementary Knowledge

	attribute
O_1	a
O_2	a
O_3	b

(1) many-valued context

	x	y	z
a	x	x	
b	x		x

(2) conceptual scale

	x	y	z
O_1	x	x	
O_2	x	x	
O_3	x		x

(3) scaled context

Utilities:

- plain scaling
- power set scaling
- relational scaling
- deduction of anonymity sets from message contents



C. Díaz, S. Seys, J. Claessens, and B. Preneel.

Towards measuring anonymity.

In R. Dingledine and P. Syverson, editors, *Proceedings of PET 2002*. Springer-Verlag, LNCS 2482, April 2002.



S. Steinbrecher and S. Köpsell.

Modelling unlinkability.

In R. Dingledine, editor, *Proceedings of PET 2003*. Springer-Verlag, LNCS 2760, March 2003.



Sebastian Clauß and Stefan Schiffner.

Structuring anonymity metrics.

In Atsuhiko Goto, editor, *Proceedings of DIM 2006*, pages 55–62, Fairfax, Virginia, USA, November 2006. ACM.



B. Ganter and R. Wille.

Formal Concept Analysis: Mathematical Foundations.

Springer-Verlag Berlin Heidelberg, 1999.

Conclusions

- Formal approach
- Hierarchy of profiles and their refinements
- Correlations \rightsquigarrow Links

Open Questions

- Transitive correlations
 - \rightsquigarrow Unlinkability
- Fuzzy concepts

Applications

- Data Track evaluation
- Data-minimization

Future Work

- Merge with Information-theoretic approaches
 - \rightsquigarrow Probabilistic approach

Thanks for your attention!