DOI: 10.2507/31st.daaam.proceedings.xxx

FORMAL CONCEPT ANALYSIS - THE POPULAR METHOD OF DATA ANALYSIS IN ENGINEERING

Lucie Sousedikova & Milan Adamek

This Publication has to be referred as: Sousedikova, L[ucie]; Adamek, M[ilan] (2020). Formal Concept Analysis – the Popular Method of Data Analysis in Engineering, Proceedings of the 31st DAAAM International Symposium, pp.xxxx-xxxx, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-xx-x, ISSN 1726-9679, Vienna, Austria

DOI: 10.2507/31st.daaam.proceedings.xxx

Abstract

This article deals with the theory, methods, and application of formal concept analysis in security engineering. Because the formal concept analysis is developed from mathematical theory, the first section gives a brief overview of mathematical foundations standing behind this powerful method of visualization table data. The research part is devoted to the application of formal concept analysis software on security components for the efficient selection of security equipment. In the end, the results of the analysis are completely described.

Keywords: Formal Concept Analysis, Formal Concepts, Concept Lattice, Attribute Implications, Security

1. Introduction

In the last years, completely automated techniques have been used in data analysis. However, when it is a little known about the data, visual analysis can be very useful. Such an effective analysis is the formal concept analysis (FCA) allowing to view non-trivial information that may not be obvious.

The formal concept analysis is an excellent method of data analysis that visualizes graphically dependencies valid in the tabular data which describe the relationship between objects and their attributes (Ganter & Wille, 1999). It interconnects three components of conceptual processing of data and knowledge:

- Discovery and reasoning with concepts and in data.
- Discovery and reasoning with dependencies in data.
- Visualization of data, concepts, and dependencies.

The integration of these three components makes FCA a powerful tool that has been applied in many areas of human activities. It should be mentioned the interesting application in the area of:

- Knowledge discovery and data mining for the purpose to extract a hierarchy of mined information from voluminous data (Lakhal and Stumme, 2005).
- Software engineering where the FCA is used for finding related artifacts from existing code (Cole, 2003).

- Security engineering for the purpose of analysis and visualization of data related to terrorist activities (Farley, 2006). [1,2]

The aim of this paper is to present FCA as an explorative and visual analysis providing the hidden non-trivial information about input data.

2. Mathematical Foundations behind FCA

FCA was introduced by Rudolf Wille in 1982 as a mathematical theory of lattices and ordered sets that were developed by Garrett Birkhoff and the other authors in the 1930s. Therefore, it is also called the method of concept lattices or the concept data analysis. This section describes mathematical foundations and conceptual structures related to FCA.

2.1. Fundamental Concepts

The basic concept of FCA is the formal context representing the input cross-table data. It is a triplet $\langle X, Y, I\rangle$ where:

- X is a non-empty set of objects and elements of X = $\{x_1, x_2, x_3, x_{4,}, x_5\}$ correspond to table rows.
- Y is a non-empty set of attributes and elements of Y = $\{y_1, y_2, y_3, y_4\}$ correspond to table columns.
- I is a binary relation between X and Y puts by the logical bivalent value 1 (object x has attribute y) or 0 (object x does not have attribute y).

I	y_1	y_2	y_3	y_4
x_1	1	1	1	1
x_2	1	0	1	1
x_3	0	1	1	1
x_4	0	1	1	1
x_5	1	0	0	0

Table 1. Formal context

From a formal context, FCA constructs object – attribute pairs known as clusters where A is a collection of all objects sharing all attributes from B, and B is the collection of all attributes shared by all objects from A. The highlighted rectangles represent all formal concepts included in Table 1.

I	y_1	y_2	**y_3**	**y_4**
x_1	1	1	**1**	**1**
x_2	1	0	**1**	**1**
x_3	0	1	**1**	**1**
x_4	0	1	**1**	**1**
x_5	1	0	0	0

I	y_1	**y_2**	**y_3**	**y_4**
x_1	1	**1**	**1**	**1**
x_2	1	0	1	1
x_3	0	**1**	**1**	**1**
x_4	0	**1**	**1**	**1**
x_5	1	0	0	0

I	**y_1**	y_2	**y_3**	**y_4**
x_1	**1**	1	**1**	**1**
x_2	**1**	0	**1**	**1**
x_3	0	1	1	1
x_4	0	1	1	1
x_5	1	0	0	0

I	**y_1**	y_2	y_3	y_4
x_1	**1**	1	1	1
x_2	**1**	0	1	1
x_3	0	1	1	1
x_4	0	1	1	1
x_5	**1**	0	0	0

Table 2. Formal concepts

In Table 2 above, there are formal concepts $\langle A_1, B_1\rangle = \langle\{x_1, x_2, x_3, x_4\}, \{y_3, y_4\}\rangle$, $\langle A_2, B_2\rangle = \langle\{x_1, x_3, x_4\}, \{y_2 y_3, y_4\}\rangle$, $\langle A_3, B_3\rangle = \langle\{x_1, x_2\}, \{y_1, y_3, y_4\}\rangle$ and $\langle A_4, B_4\rangle = \langle\{x_1, x_2, x_5\}, \{y_1\}\rangle$. [3]

2.2. Concept-forming Operators

All formal concepts are ordered by subconcept-superconcept relation based on inclusion relation on objects and attributes. For formal concepts $\langle A_1, B_1\rangle$ and $\langle A_2, B_2\rangle$ subconcept-superconcept ordering is defined as $\langle A_1, B_1\rangle \leq \langle A_2, B_2\rangle$ if $A_1 \subseteq A_2$ (if $B_2 \subseteq B_1$). It means $\langle A_1, B_1\rangle$ is more specific than $\langle A_2, B_2\rangle$ and $\langle A_2, B_2\rangle$ is more general than $\langle A_1, B_1\rangle$. [4,5]

Then, the previous formal concepts can be ordered as follows: $\langle A_3, B_3 \rangle \leq \langle A_1, B_1 \rangle$, $\langle A_3, B_3 \rangle \leq \langle A_2, B_2 \rangle$, $\langle A_3, B_3 \rangle \leq \langle A_4, B_4 \rangle$, $\langle A_1, B_1 \rangle \parallel \langle A_4, B_4 \rangle$ (incomporable), $\langle A_2, B_2 \rangle \parallel \langle A_4, B_4 \rangle$.

All ordered formal concepts of a given formal context create a concept lattice which includes all potentially interesting concepts in data.

3. Application of FCA in Security Engineering

Currently, there is a wide range of security features on the market. The security industry is constantly coming up with new and modern components in whose technical parameters the average user gets lost.

FCA is the ideal tool for the efficient selection of an appropriate component with the required level of security. As a method of exploratory analysis of object-attribute tabular data allows the easy orientation in requirements for security systems. It provides a comprehensive overview of security components (objects) and their technical parameters (attributes) due to the graphical visualization of dependencies contained in the data. Then the analysis of the available security equipment can be used to optimize the security design in the projecting of the security system.

There is a large number of software tools used for FCA, e.g. the Concept Explorer which was used in this research and which allows to:

- Context editing.
- Create a conceptual lattice.
- Find implications or association rules valid in context.
- Conduct an attribute survey.

Concept Explorer creates conceptual lattices from tabular data where objects are situated in table rows, and attributes are in columns. If the object has the given attribute, the box in the table is marked with a bivalent value 1, and if not, the box is marked with a bivalent value 0. The next table depicts the tabular data with objects (laminated safety glass Stratobel Security Burglary marked as SSB) and their attributes (resistance class, total thickness, weight, acoustic isolation, and dimension) which are processed by the program as follows.

Stratobel Security Bullet (SSB)	Resistance class according to norms																				Total thickness			Weight			Acou-stics		Dimen-sion		Tests with Kalashnikov weapon or equivalent
	Break-in EN 356					Bullets EN 1063																									
	P3A-P4A	P5A	P6B	P7B	P8B	BR1 S	BR1 NS	BR2 S	BR2 NS	BR3 S	BR3 NS	BR4 S	BR4 NS	BR5 S	BR5 NS	BR6 S	BR6 NS	BR7 NS	SG1 S	SG2 S	14-34 mm	35-55 mm	56-80 mm	30-85 kg/m2	86-140 kg/m2	141-190 kg/m2	37-42 dB	43-47 dB	600x260 cm	600x321 cm	
SSB 402-1-B	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	1	0
SSB 802-5-B	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	1	0
SSB 902-7-B	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	1	0
SSB 104-1-B	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	0	0
SSB 403-5-B	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	1	0
SSB 704-3-B	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	1	0	0
SSB 004-8-B	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	1	0	0
SSB 304-6-B	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	1	1	0	0
SSB 106-1-B	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	1	0	0
SSB 504-4-B	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	1	0	0	0	1	1	0	1
SSB 806-2-B	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	1	1	0	1
SSB 905-9-B	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	1	1	0	1
SSB 408-1-B	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	1	1	0	0
SSB 009-1-B	0	0	0	[illegible]	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	1	1	0	0

Table 3. Laminated safety glass Stratobel Security Bullet with technical parameters

From such a context table, two kinds of output are generated by Concept Explorer software. The first is a graphical image of the concept lattice in which the objects and attributes are interconnected according to how they belong to each other. It is a collection of formal concepts included in input data and hierarchically arranged according to their subconcept-superconcept relation.

In the following figure is depicted the conceptual lattice where the highest point represents the set of all objects, the lowest point represents the set of all attributes, and the individual points represent:

- Objects (white rectangles).
- Attributes (grey rectangles).

- Connection to an object (black semicircle).
- Connection to an attribute (blue semicircle).

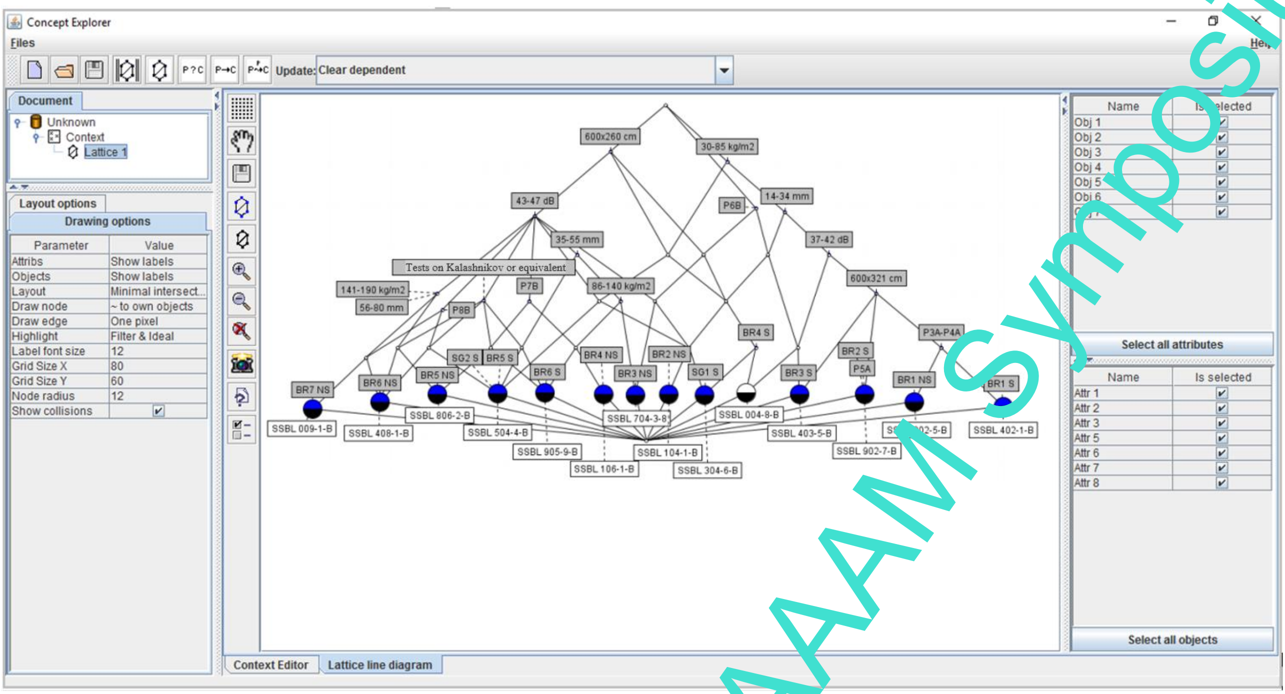

Fig. 1. Concept lattice of SSB glass in the environment of Concept Explorer software

The created conceptual lattice can be modified in various ways. It is possible to form it into various forms, to manipulate the names of its objects and attributes, or to set the different sizes of its nodes and edges. Also, we can display only the required values.

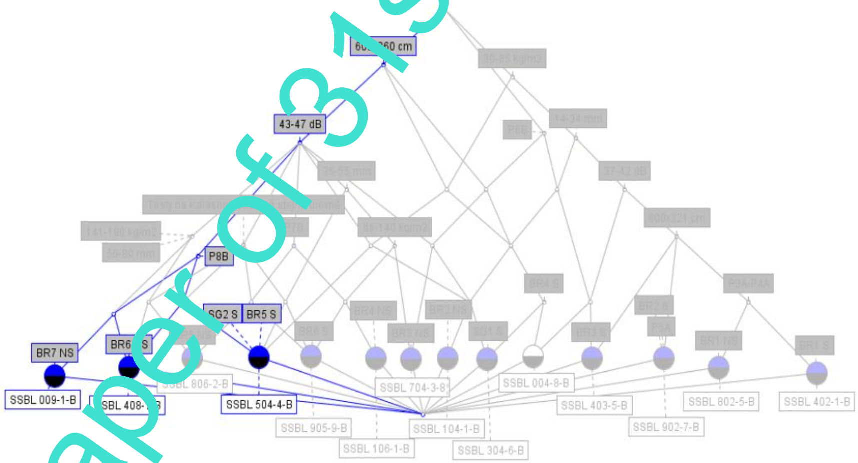

Fig. 2. Display of the required parameter P8B - navigating in concept lattice [6]

The second output of FCA is a collection of attribute implications which semantically follow from others and describe a particular dependency which are true in the input data. [7]

1 < 2 > P3A-P4A ==> 14-34 mm 30-85 kg/m2 37-42 dB 600x321 cm;
2 < 1 > P5A ==> BR2 S 14-34 mm 30-85 kg/m2 37-42 dB 600x321 cm;
3 < 3 > P7B ==> 43-47 dB 600x260 cm;
4 < 3 > P8B ==> 43-47 dB 600x260 cm;
5 < 1 > BR1 S ==> P3A-P4A 14-34 mm 30-85 kg/m2 37-42 dB 600x321 cm;
6 < 1 > BR1 NS ==> P3A-P4A 14-34 mm 30-85 kg/m2 37-42 dB 600x321 cm;
7 < 1 > BR2 S ==> P5A 14-34 mm 30-85 kg/m2 37-42 dB 600x321 cm;
8 < 1 > BR2 NS ==> P6B 14-34 mm 30-85 kg/m2 43-47 dB 600x260 cm;
9 < 1 > BR3 S ==> P6B 14-34 mm 30-85 kg/m2 37-42 dB 600x321 cm;
10 < 1 > BR3 NS ==> P6B 35-55 mm 86-140 kg/m2 43-47 dB 600x260 cm;

Fig. 3.Example of attribute implications

5. Conclusion

The research suggests that FCA is the perfect technique, as a method of visualization of hidden dependencies valid in the data, how to select effective security equipment. In this paper, it was described the application of FCA to security engineering for analysis of tabular data with security components and their technical parameters. Concretely, it was analysed laminated safety glass Stratobel Security Bullet which has different characteristics and class resistance against break-in and bullets. The object-attribute input data were processed in Concept Explorer software based on FCA, and two types of graphical outputs were computed – concept lattice and attribute implications. Both results lead to a better and straightforward interpretation of tabular data and to the efficient possibility of how to choose an appropriate security element with the required level of resistance. FCA is a method which deals with object-attribute data set within a data table with bivalent logical values. That is why the objects and attributes were converted directly to the binary tables. It should be noted that this is the main limitation of FCA – preprocessing and conversion of attribute values into a set of binary attributes. Moreover, if the attributes are fuzzy, it is not possible to realize analysis without conceptual scaling. [8] This issue should be discussed in a future research paper by fuzzy logic describing fuzzy attributes that are more in common in the real world.

6. Acknowledgment

This work was supported by the Internal Grant Agency of Tomas Bata University under the project No. IGA/CebiaTech/2020/003.

7. References

[1] Ganter, B.; Wille, R. (1999). Formal Concept Analysis; Springer-Verlag Berlin Heidelberg, ISBN 978-3-540-62771-5, Berlin.
[2] Belohlavek, R. (2008). Introduction to Formal Concept Analysis, Available from: http://belohlavek.inf.upol.cz/vyuka/IntroFCA.pdf Accessed: 2020-07-01
[3] Belohlavek, R.; Vychodil, V. (2005). What is a fuzzy concept lattice?, Proceedings of the 3rd international conference on Concept Lattices and their Applications, ISBN 80–248–0863–3, Snasel, V. (Ed.), pp. 34-45, Olomouc, Czech Republic.
[4] Martínez Morales, M.; Alatriste Trujillo A. & Guerra-Hernández A. (2004). Concept lattice learning in analytic induction, Available from: https://www.researchgate.net/publication/236373258_Concept_lattice_learning_in_analytic_induction Accessed: 2020-08-08
[5] Belohlavek, R.; Kostak, M. & Osicka, P. (2013). Formal concept analysis with background knowledge: a case study in paleobiological taxonomy of belemnites, International Journal of General Systems. Vol. 42, No. 4, 2013, pp. 426-440, ISSN 0308-1079, DOI http://dx.doi.org/10.1080/03081079.2013.765079
[6] Sousedikova, L. (2018). The Application of Port-Royal Logic for the Selected Plastic Casing Protection Features of Sites, Thesis, Faculty of Applied Informatics, Tomas Bata University in Zlin, Zlin, Czech Republic
[7] Sousedikova, L. (2020). Application of Formal Concept Analysis in Security Engineering, Trilobit, 2020, pp. 1-6, ISSN 1804-1795
[8] Butka, P. (2012). Use of Concept Lattices for Data Tables with Different Types of Attributes. Journal of Information and Organizational Sciences, Vol. 36, No. 1, 2012, pp. 1-12, ISSN 1846-3312
[9] Lakhal, L. & Stumme, G. (2005). Efficient mining of association rules based on formal concept analysis. In: Formal Concept Analysis. Lecture Notes in Computer Science, Ganter, B.; Stumme, G.; Wille, R. (Eds.), pp 180-195, Springer, ISBN 978-3-540-27891-7, Berlin, Heidelberg
[10] Cole, R. J.; Tilley T. (2003). Conceptual analysis of software structure, Available from: https://www.researchgate.net/publication/2890916_Conceptual_Analysis_of_Software_Structure Accessed: 2020-08-08
[11] Farley, J. D. (2006). https://www.nytimes.com/2006/05/16/opinion/16farley.html, Available from: http://www.exact-address-of-site Accessed: 2020-08-08