Clones in Social Networks

Stephan Doerfel[1], Tom Hanika[2,3], and Gerd Stumme[2,3]

[1] Micromata GmbH Kassel, Germany
[2] Knowledge & Data Engineering Group
University of Kassel, Germany

[3] Interdisciplinary Research Center for Information System Design
University of Kassel, Germany

sdo@cs.uni-kassel.de, tom.hanika@cs.uni-kassel.de,
stumme@cs.uni-kassel.de

Abstract It is well known that any bipartite (social) network can be regarded as a formal context (G, M, I). Therefore, such networks give raise to formal concept lattices which can be investigated utilizing the toolset of Formal Concept Analysis (FCA). In particular, the notion of clones in closure systems on M, i. e., pairwise interchangeable attributes that leave the closure system unchanged, suggests itself naturally as a candidate to be analyzed in the realm of FCA based social network analysis. In this study, we investigate the notion of clones in social networks. After building up some theoretical background for the clone relation in formal contexts we try to find clones in real word data sets. To this end, we provide an experimental evaluation on nine mostly well known social networks and provide some first insights on the impact of clones. We conclude our work by nourishing the understanding of clones by generalizing those to permutations of higher order.

Keywords: Social Network Analysis, Formal Concept Analysis, Clones

1 Introduction

Clone items in a family of sets $\mathcal{F}$ on some set M are pairs of elements of M that leave $\mathcal{F}$ invariant when being interchanged in every element of $\mathcal{F}$. The notion of clones was initially proposed[4] in "Clone items: a pre-processing information for knowledge discovery" by R. Medina and L. Nourine. The structural idea is tempting and has become a well considered notion in formal concept analysis (FCA) since then. The authors of [7] unraveled some of the "hidden combinatorics" by showing possible concept lattice factorizations. Even very simple formal contexts may have an ample amount of clones. Those observations led to computational

The authors of this work are given in alphabetical order. No priority in authorship is implied.

[4] This work is noted to be submitted (e. g, in [7]), but has never been published.

investigations like [14]. Also theoretical applications for association rule mining were made [13].

All this strikes the question for applications in real world data. The first (and only) attempt the authors are aware of was made in [11]. There, three different data sets, in particular *Mushroom, Adults,* and *Anonymous* from the UCI Machine Learning Repository [9], were examined. The observations were sobering, in fact, two data sets were free of clones whereas the mushroom data set had only few. However, the main question – are clone items frequent in natural data sets, in particular in social network data – is not answered in general.

Taking advantage of the strong correspondence between bipartite social networks and formal contexts, we investigate nine well-known social networks of different sizes for clone items. To complement this we consider two non-social network data sets as well. Building on this, we present insights on why clones are not common in social networks using a characterization of clones on the level of the formal context. We conclude that clones are not suitable subjects of investigation in social network data.

To resolve this dilemma, we point out a more general notion of clones. For this we fall back to permutations on the set of attributes in a formal context, providing a natural extension of the clone property. These *higher order clones* are able to identify more complicated "clone structures" and should be the next step in the investigation of relational data structures.

This work is structured as follows. In Section 2, we recall basic notations of FCA, introduce social networks, and show the correspondence between them. Then, in Section 3, we provide a characterization of clone items on the level of formal contexts. Following this, in Section 4 we demonstrate how the notion of clones can be applied in the realm of social networks. Subsequent to experiments on various data sets, in Section 5, we extend the notion of clone items to higher order clones. Eventually, we conclude our work with Section 6.

2 Preliminaries

We give a short recollection of the ideas from formal concept analysis as introduced in [5, 18] that are relevant in this work. We use the common presentation of formal contexts by $\mathbb{K} = (G, M, I)$, where G and M are sets and $I \subseteq G \times M$. The elements of G are called objects, those of M are called attributes, and $(g, m) \in I$ signifies that object g has the attribute m. The correspondence to a bipartite graph (network) is at hand. Let $H = (U \cup W, E)$ be such an undirected bipartite graph with $U \cap W = \emptyset$ where U is the set of individuals (often users), W some set of common properties, and $E \subseteq \{\{u, w\} \mid u \in U, w \in W\}$ the set of edges between U and W. There are two natural ways of identifying H as a formal context. In the following, we choose $\mathbb{K}(H) = (U, W, I)$ as the to H associated formal context,[5] where for $u \in U$ and $w \in W$, we have $(u, w) \in I \Leftrightarrow \exists e \in E : u \in e \wedge w \in e$. In the following we use the terms network, bipartite graph, and formal context interchangeably in the sense above.

[5] The second way yields the dual context $\mathbb{K}(H) = (W, U, I)$.

We will utilize the common *derivation* operators $\cdot'\colon \mathcal{P}(G) \to \mathcal{P}(M), A \mapsto B \coloneqq \{b \in M \mid \forall g \in A\colon (g,m) \in I\}$ and $\cdot'\colon \mathcal{P}(M) \to \mathcal{P}(G), B \mapsto A \coloneqq \{a \in G \mid \forall m \in B\colon (g,m) \in I\}$. Having those operations we call a formal context $\mathbb{K} = (G,M,I)$ *object clarified* iff $\forall g,h \in G, g \neq h : g' \neq h'$, *attribute clarified* iff $\forall m,n \in M, m \neq n : m' \neq n'$ and *clarified* iff it is both. In this definition we used g' as shorthand for $\{g\}'$. Clarification will later on correspond to a particular trivial kind of clones. Similarly we call a clarified context $\mathbb{K}$ *object reduced* if for all $g \in G$ there is no $S \subseteq G \setminus \{g\}$ such that $g' = S'$. We call $\mathbb{K}$ *attribute reduced* iff for all $m \in M$ there is no $S \subseteq M \setminus \{m\}$ such that $m' = S'$. And, we call this $\mathbb{K}$ *reduced* iff $\mathbb{K}$ is attribute and object reduced.

A pair (A,B) where $A \subseteq G$, $B \subseteq M$ with $A' = B$ and $B' = A$ is called a *formal concept*. Here, A is called the *concept extent* and B is called the *concept intent*. The set of all these formal concepts, i.e., $\mathfrak{B}(\mathbb{K}) \coloneqq \{(A,B) \mid A \subseteq G, B \subseteq M, A' = B, B' = A\}$ gives raise to an order structure $(\mathfrak{B}, \leq)$ using $(A,B) \leq (C,D)\colon \Leftrightarrow A \subseteq C$, called *concept lattice*. For clone items we are particularly interested in the two entailed closure system, i.e, in the *object closure system* $\mathfrak{G}(\mathbb{K}) \coloneqq \{A \in G \mid (A,B) \in \mathfrak{B}(\mathbb{K})\}$ and the *attribute closure system* $\mathfrak{M}(\mathbb{K}) \coloneqq \{B \in M \mid (A,B) \in \mathfrak{B}(\mathbb{K})\}$. We may denote those by $\mathfrak{G}$ and $\mathfrak{M}$ whenever the according context is implicitly given.

Clones Besides the original definition of what clone items are there will be some graduations useful to social networks. We start with the common definition. Given a formal context $\mathbb{K} = (G,M,I)$ and two items $a, b \in M$, we say *a is clone to b* in $\mathfrak{M}$ if $\forall X \in \mathfrak{M}\colon \varphi_{a,b}(X) \in \mathfrak{M}$, with:

$$\varphi_{a,b}(X) \coloneqq \begin{cases} X \setminus \{a\} \cup \{b\} & \text{if } a \in X \wedge b \notin X \\ X \setminus \{b\} \cup \{a\} & \text{if } a \notin X \wedge b \in X \\ X & \text{else} \end{cases}$$

We may denote this property by $a \sim_{\mathbb{K}} b$ and whenever the context is distinctive $a \sim b$. It is obvious that $\sim$ is a reflexive and symmetric relation on $M \times M$. Actually, it is also transitive, which can be shown easily, hence $\sim$ is an equivalence relation. Since every $a \in M$ is a clone to itself we say an a is a *proper clone* iff there is a $b \in M \setminus \{a\}$ such that $a \sim b$. In a not-clarified formal context there might be some $m, n \in M, m \neq n$ such that $m' = n'$. Those elements are proper clones. However, this is obvious and not revealing any hidden structure besides the fact that two identical copies are present. Therefore we call a proper clone $a \in M$ *trivial* iff there is a $b \in M \setminus \{a\}$ with $a' = b'$.

A this point one may ask if it is hard to construct a formal context having a significant number of clones. Actually, this is very easy as the following example discloses.

Example 2.1. The nominal scales, i.e., $(\{1,\ldots,n\},\{1,\ldots,n\},=)$ and the contranominal-scale $(\{1,\ldots,n\},\{1,\ldots,n\},\neq)$ provide formal contexts where every attribute element is a non-trivial clone, see 1. Furthermore, the union of two formal contexts, i.e., $\mathbb{K}_1 \coloneqq (G_1,M_1,I_1)$ and $\mathbb{K} \coloneqq (G_2,M_2,I_2)$ becomes $\mathbb{K}_1 \cup \mathbb{K}_2 \coloneqq (G_1 \cup G_2, M_1 \cup M_2, I_1 \cup I_2)$, preserves the clones from $\mathbb{K}_1$ and $\mathbb{K}_2$.

N-scale	m_1	m_2	m_3
g_1	×		
g_2		×	
g_3			×

CN-scale	m_1	m_2	m_3
g_1		×	×
g_2	×		×
g_3	×	×	

Figure 1. Small example contexts with all attribute elements being clone to each other. In particular the nominal-scale (left) and the contra-nominal-scale (right).

All the above can be defined similarly for elements of G using the dual-context, i.e., the context where objects and attributes are interchanged. We therefore omit the explicit definitions and continue assuming the necessary definitions are made. However, we may provide some wording to differentiate between clones in $\mathfrak{M}$ and clones in $\mathfrak{G}$ for some formal context (G, M, I). When necessary we call the former *attribute clone* and the latter *object clone.*

3 Theoretical observations

In this section, we derive some useful properties of clones as well as a characterization of the clone propery on the level of the context table.

Lemma 3.1 (Clones are incomparable)**.** *Let $\mathbb{K} = (G, M, I)$ be a formal context and $a, b \in M$. If $a \sim b$, then from $a' \subseteq b'$ follows $a' = b'$.*

Proof. Using $a' \subseteq b'$ we show $b' \subseteq a'$. We examine the mapping

$$\varphi_{ab}(b'') = \begin{cases} b'' & \text{if } a \in b'' \\ b'' \setminus \{b\} \cup \{a\} & \text{if } a \notin b''. \end{cases}$$

We show that the second case is invalid. From $a \sim b$ and $\varphi_{ab}(b'')$ being a closure we deduce $a'' \subseteq b'' \setminus \{b\} \cup \{a\}$. Since $a' \subseteq b'$, we have $b'' \subseteq a''$ and together we yield $b \in b'' \subseteq a'' \subseteq b'' \setminus \{b\} \cup \{a\}$ contradicting the case. Hence, only the first case can exist, meaning $a \in b''$, thus obviously $b' \subseteq a'$. □

Lemma 3.2 (Clone irreducability)**.** *Let $\mathbb{K} = (G, M, I)$ be a clarified formal context and attributes $a, b \in M : a \neq b$ with $a \sim b$. Then a is irreducible in $\mathbb{K}$.*

Proof. Assume a is reducible, i.e., there exists a set of attributes $N \subseteq M$ with $a \notin N$ and $\bigcap_{n \in N} n' = a'$. As $\mathbb{K}$ is clarified, we have $a' \neq b'$, thus from Lemma 3.1 follows $b \notin a''$. Therefore $\varphi_{a,b}(a'') = a'' \setminus \{a\} \cup \{b\}$. From the reducibility assumption follows

$$\forall n \in N : n' \supseteq a' \Longrightarrow n \in a'' \overset{n \neq a}{\Longrightarrow} n \in a'' \setminus \{a\} \cup \{b\} = \varphi_{a,b}(a'').$$

Thus, $a' = \bigcap_{n \in N} n' \supseteq \varphi_{a,b}(a'')'$, which means $a'' \subseteq \varphi_{a,b}(a'') = a'' \setminus \{a\} \cup \{b\}$. Clearly, this means $a = b$ contradicting the lemma's assumption. □

While clarifying a context removes the non-trivial clones, additionally reducing that context does not change the clone relationship any further. Therefore, for finding non-trivial clones it suffices considering reduced contexts. Next, we describe for such contexts how clones can be identified directly from the context's table. [7] already found it is sufficient to check join-irreducible intents to check the clone property. The respective result there (Proposition 1) is formulated for the dual version of formal contexts, i. e. where G and M are interchanged. Also, for the proof the authors of [7] refer to a manuscript that had been submitted (at the time) but appears to have never been published. For the sake of completeness, we present a variation of their result in the common notion of a formal context and present a proof. Here, we already use the fact that in a reduced context, the join irreducible concepts are exactly the object concepts.

Theorem 3.1. *Let $\mathbb{K} = (G, M, I)$ be a reduced formal context and $a, b \in M$ with $a \neq b$. The following are equivalent:*

1. *$a \sim b$*
2. *For each object $g \in G$, there is an object $h \in G$ such that $\varphi_{a,b}(g') = h'$.*

Proof. First we show, 1. $\Longrightarrow$ 2. For $a, b \in g'$ or $a, b \notin g'$, the claim is obvious (using $h := g$). Without loss of generality, we can assume $a \in g'$ and $b \notin g'$, thus $\varphi_{a,b}(g') = g' \setminus \{a\} \cup \{b\}$.

As $\varphi_{a,b}(g')$ is an intent, there exists a set of attributes $H \subseteq G$ with $H' = \varphi_{a,b}(g') = g' \setminus \{a\} \cup \{b\}$. In particular, this means $b' \supseteq H$ and therefore, we can partition H into $H_{a,b} := H \cap G_{a,b}$ and $H_{\bar{a},b} := H \cap G_{\bar{a},b}$. As clearly $a \notin \varphi_{a,b}(g')$, $H_{\bar{a},b}$ cannot be empty. We yield:

$$\begin{aligned}
& & g' \setminus \{a\} \cup b &= \bigcap_{h \in H_{a,b}} h' \cap \bigcap_{h \in H_{\bar{a},b}} h' \\
H_{\bar{a},b} \neq \emptyset &\Longrightarrow & (g' \setminus \{a\} \cup b) \setminus \{b\} &= \bigcap_{h \in H_{a,b}} h' \cap \bigcap_{h \in H_{\bar{a},b}} (h' \setminus \{b\}) \\
b \notin g' &\Longrightarrow & g' \setminus \{a\} &= \bigcap_{h \in H_{a,b}} h' \cap \bigcap_{h \in H_{\bar{a},b}} (h' \setminus \{b\}) \\
H_{a,b} \subseteq G_{a,b} &\Longrightarrow & g' &= \bigcap_{h \in H_{a,b}} h' \cap \bigcap_{h \in H_{\bar{a},b}} (h' \setminus \{b\} \cup \{a\}) \\
H_{\bar{a},b} \subseteq G_{\bar{a},b} &\Longrightarrow & g' &= \bigcap_{h \in H_{a,b}} h' \cap \bigcap_{h \in H_{\bar{a},b}} \varphi_{a,b}(h')
\end{aligned}$$

As g is irreducible, we either have an object $h \in H_{a,b}$ with $g' = h'$ or an object $h \in H_{\bar{a},b}$ with $g' = \varphi_{a,b}(h')$. Clearly, the former cannot be true, as $b \in h'$ for $h \in H_{a,b}$ and $b \notin g'$. From the latter follows $\varphi_{a,b}(g') = h'$).

Next, we show 2. $\Longrightarrow$ 1 : Let $N \subseteq M$ be an intent of $\mathbb{K}$, i. e. there is a set of attributes $H \subseteq G$ such that $N = H'$. We show that $\varphi_{a,b}(N)$ is an intent. This is trivial for the cases $a, b \in N$ and $a, b \notin N$. Without loss of generality, we assume

$a \in N$ and $b \notin N$. Then

$$\varphi_{a,b}(N) = H' \setminus \{a\} \cup \{b\} = \bigcap_{h \in H} (h' \setminus \{a\} \cup \{b\})\,.$$

Since $a \in N$, for each $h \in H$ the set $h' \setminus \{a\} \cup \{b\} = \varphi_{a,b}(h')$ is an intent of $\mathbb{K}$ and therefore $\varphi_{a,b}(N)$ is an intent as well. □

The theorem characterizes clones on the context level: Two attributes a and b are clones if for each object $g \in G$ whose row contains only one of the two attributes, there is another object $h \in G$ such that its row contains only the other of the two attributes, while the remaining parts of the rows are identical, i. e., $g' \setminus \{a\} = h' \setminus \{b\}$.

4 Clones in social networks

In the following, we identify any given bipartite social network $H = (U \cup W, E)$ with the formal context counterpart $\mathbb{K} = (U, W, I)$, with I as introduced in Section 2. Furthermore, we may also identify any single-mode social network (U, E) with (U, U, I) in the natural way. Transferring the definitions from Section 2, we obtain what clones in social networks are. In particular, object clones can be identified as *user clones* and attribute clones are either some *property clone*, in the bipartite case, or also user clones, in the single-mode case.

The main tempting question now is: are there natural occurring user and property clones? We try to answer this in the following section.

4.1 Experiments

Data set description Almost all of the following data sets can be obtained from the UCI Machine Learning Repository [9]. We consider nine social network graphs and two non social network data sets.

zoo: Obtained form [9] and contributed by Richard Forsyth, a data set consisting of 101 animals and seventeen attributes, fifteen Boolean and two numerical. The Boolean were nominal scaled as well as the numerical values. Hence, having a property and not having a property are two different attributes. This results in an attribute set with 43 elements.

cancer: A data set on breast cancer diagnosis [12] consisting of 699 instances and ten numerical attributes, which were nominal scaled.

facebooklike: Obtained and found via [15]. This data set emerged form an online community of students from the University of California, Irvine. In a forum 337 users were communicating via 522 topics.

southern: This classical small world social network consists of fourteen woman attending eighteen different events, well investigated in [17].

club: A network consisting of 25 corporate executive officers and fifteen social clubs in which they are involved in [3].

movies: A set of 39 composers of Hollywood film music were related to 62 producers [4].

aplnm: Social network linking 79 participants from the *Lange Nacht der Musik* of 2013, to 188 events they participated in, in depth investigated in [1].

jazz: Collaboration network of jazz musicians obtained from [6].

dolphin: In this data set nodes are bottlenose dolphins of a community living off Doubtful Sound, a fjord in New Zealand. An edge between two dolphins indicates a frequent contact [10].

hightech: Some social network from [16] within the parameters of a social network but with no further insights provided.

wiki: This data set contains all the Wikipedia voting data [8, 16] starting from 2001 until 2008. Nodes represent voters and to be voted on Wikipedia users, which we partitioned.

For comparison, we also investigate randomized versions of all those data sets, generated using a coin draw process. This may imply that the resulting formal contexts are prone to the stegosaurus phenomenon. However, no unbiased method for generating formal context for a given number of objects, attributes, and density is known [2].

Computation Computing the attribute (object) clones for a given formal context (G, M, I) would imply to know the associated attribute (object) closure system. However, computing those is computational infeasible for contexts of a particular size or greater. To cope with this barrier we utilize Lemma 3.2 and Theorem 3.1. Hence, instead of checking all elements of a closure system we only need to check the irreducibles. Therefore we checked brute force all combinations of attributes (objects) for every given formal context (social network) by checking the according irreducibles.

In particular we computed for every clone the number of tivial and non-trivial object clones, and attribute clones. The results are shown in Table 1. In addition we also computed the number of trivial and non-trivial clones for the object/attribute-projections for every formal context. For given formal context (G, M, I) the object projection is a new formal context (G, G, I_G) with $(g, h) \in I_G \Leftrightarrow \exists m \in M : (g, m) \in I \wedge (h, m) \in I$. The attribute projection can be obtained on a similar way. Those so-called one-mode projections are often used in social network analysis. The computational results for those contexts are shown in Table 2.

Discussion The most obvious result for all data sets alike is that non-trivial clones are very infrequent. Omitting the wiki data set only two data sets have clones at all, in particular a very small number of object clones compared to the size of the network. We investigated the exception by the wiki data set further and discovered a large nominal scale as subcontext responsible for the vast amount of clones. Since the wiki data set is the result of a collection of voting processes this would represent single votes. For trivial clones we have

Table 1. Properties of the considered (social) networks and data sets and results for clone experiment. The suffix "-r" denotes the randomized version of a given data set. With tG we denote trivial clones whereas clones denote non-trivial clones.

Name	$\|U\|$	$\|M\|$	density	# G-clones	# M-clones	# tG-clones	# tM-Clones
zoo	101	43	0.390	0	0	42	2
zoo-r	101	43	0.385	0	0	0	0
cancer	699	92	0.110	0	0	236	0
cancer-r	699	92	0.106	0	0	0	0
facebooklike	377	522	0.014	7	0	24	83
facebooklike-r	377	522	0.014	0	0	0	0
southern	18	14	0.352	0	0	1	1
southern-r	18	14	0.309	0	0	0	0
aplnm	79	188	0.061	0	0	1	21
aplnm-r	79	188	0.056	0	0	0	0
club	25	15	0.250	0	0	0	0
club-r	25	15	0.261	0	0	0	0
movies	62	39	0.079	0	0	1	0
movies-r	62	39	0.074	0	0	5	0
jazz	198	198	0.068	7	0	0	0
jazz-r	198	198	0.068	0	0	0	0
dolphins	62	62	0.082	0	0	2	2
dolphins-r	62	62	0.058	0	0	4	3
hightech	33	33	0.148	0	0	1	1
hightech-r	33	33	0.149	0	0	0	0
wiki	764	605	0.006	234	234	73	30
wiki-r	764	605	0.006	0	0	21	2

diverse observations. Some networks like facebooklike have a significant amount of trivial clones. Others however of comparable size do not, like jazz. Since those clones do not reveal any hidden structure but the fact that copies of users or properties are present in the network, we consider these clones uninteresting.

For the object and attribute projections we obtain almost the same results. Almost no non-trivial clones are present. Though, the number of trivial clones has increased in almost all the networks. This could be another revelation that simple one-mode projections are insufficient for analyzing bipartite networks.

All in all, the notion of non-trivial clones seems insufficient for the investigation of social networks. This is not surprising while recalling the theoretical results from Section 3. The structural conditions for obtaining a clone are to strong. However, it strikes the question if there is a generalization which is softening those conditions while preserving enough structure.

Table 2. Properties of the projected (social) networks and data sets. V denotes the set of vertices. The suffix π_G represents a projection on G whereas π_M denotes a projection on M.

Name	$\|V\|$	edge density	# clone-pairs	# trivial clones
zoo:π_G	101	1	0	100
zoo:π_M	43	0.782	0	2
cancer:π_G	699	0.860	0	296
cancer:π_M	92	0.660	0	0
facebooklike:π_G	377	0.015	7	25
facebooklike:π_M	522	0.114	6	87
southern:π_G	18	0.913	0	13
southern:π_M	14	0.744	0	11
aplnm:π_G	79	0.442	0	1
aplnm:π_M	188	0.390	0	21
club:π_G	25	0.870	2	16
club:π_M	15	0.653	0	0
movies:π_G	62	0.288	0	1
movies:π_M	39	0.313	0	0
jazz:π_G	198	0.680	0	27
jazz:π_M	198	0.680	0	27
dolphins:π_G	62	0.312	0	2
dolphins:π_M	62	0.312	0	2
hightech:π_G	33	0.592	0	1
hightech:π_M	33	0.592	0	1
wiki:π_G	764	0.056	255	96
wiki:π_M	605	0.040	232	43

5 Generalized Clones

The results from the previous section motivate finding a more general clone notion for formal contexts. In [7] the authors provided an interesting generalization of clones in a formal context. They proposed P-Clones, i. e., clones with respect to the family of pseudo intents, and A-Clones, i. e., clones in a particular kind of atomized context. Both approaches are based on using some kind of modified family of sets. Another course of action was taken in [11], in which the author used a measure of "cloneniness" based on the number of incorrect mapped sets.

We take a different approach, using the original set of closures – the intents – based on the following observation.

Remark 5.1 (Clone permutation). Every pair (a, b) of elements $a, b \in M$ with $a \sim b$ for a given formal context (G, M, I) gives raise to a permutation $\sigma : M \to$

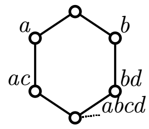

Figure 2. Example of a clone-free closure system on four attributes.

$M, m \mapsto \sigma(m)$, with $\sigma(a) = b, \sigma(b) = a$, and $\sigma(m) = m$ for $m \in M \setminus \{a, b\}$. We denote such permutations as *clone permutations.*

Since for every $a \in M$ we have $a \sim a$, the set of clone permutations S for a given formal context (G, M, I) contains the identity. For any two elements $a, b \in M$ with $a \sim b$ we can represent the assocation clone permutation σ by $\sigma := (ab)$ using the reduced cycle notation. From this we note that the set of all pairs of proper clones corresponds to a particular subset of permutations on M where every permutation σ contains exactly one two-cycle. This gives rise to two possible generalizations.

5.1 Multiple two-cycles

We motivate this approach using the lattice for a closure system on $M = \{a, b, c, d\}$ represented in Figure 2. In this closure system there are no proper clones. However, we can find a permutation σ that preserves the closure system. For example, the permutation $\sigma = (ab)(cd)$, which is a permutation of two disjoint cycles of length two. This permutation is not representable by exactly one cycle of length two.

Hence, we propose permutations representable as products of cycles of length two as one generalization of clones. Yet, this immediately gives raise to the idea of higher order permutations.

5.2 Higher Order

Again, we want to motivate this generalization by providing an example. In Figure 3 (left), we show the lattice for a closure system $M = \{a, b, c, d\}$. This closure system is free of (proper) clones. However, we find a permutation $\sigma = (ab)(cd)$ in the above described manner. In addition we find a permutation of order four, i. e., $\sigma^4 = id$, preserving the closure system, e. g., $\sigma = (acbd)$. In the same figure on the right we observe a permutation of order five, i.e., $\sigma = (acedb)$, answering the natural question for an permutation with odd order.

6 Conclusion

While starting the investigation the authors of this work were confident to discover clones in social networks, at least for networks of a particular minimal

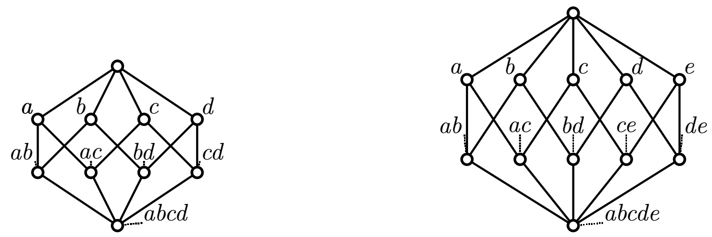

Figure 3. Example for a clone-free closure system on four attributes (left) and on five attributes (right).

size. In order to cope with the computational complexity of closure systems we utilized results from [7] and expressed them in terms of statements about formal contexts. However, our investigation did not reveal any meaningful connection between clones and social networks. The only significant observation was the emergence of trivial clones while projecting bipartite social networks to one set of nodes.

This setback, though, led us to discover two more general notion of clones, which can cope with more structural requirements. Investigating those more thoroughly should be the next step in clone related research. To this end, we finish our work with the following three open questions.

Question 1: To which graph theoretical notion could the idea of clone permutation correspond to?

Question 2: Does the set of all valid clone permutations on a closure set always form a group and if no, why not?

Question 3: If yes, can this group provide new insights into the structure of closure systems or of social networks?

References

[1] D. Borchmann and T. Hanika. “Individuality in Social Networks.” In: *Formal Concept Analysis of Social Networks*. Ed. by Rokia Missaoui, Sergei O. Kuznetsov, and Sergei Obiedkov. Cham: Springer International Publishing, 2017, pp. 19–40.

[2] Daniel Borchmann and Tom Hanika. “Some Experimental Results on Randomly Generating Formal Contexts.” In: *CLA*. Ed. by Marianne Huchard and Sergei Kuznetsov. Vol. 1624. CEUR Workshop Proceedings. CEUR-WS.org, 2016, pp. 57–69.

[3] *Club membership network dataset – KONECT*. Jan. 2016.

[4] R.R. Faulkner. *Music on Demand: Composers and Careers in the Hollywood Film Industry*. Composers and Careers in the Hollywood Film Industry. Transaction Publishers, 2003.

[5] B. Ganter and R. Wille. *Formal Concept Analysis: Mathematical Foundations.* Springer-Verlag, Berlin, 1999, pp. x+284.

[6] Pablo Gleiser and Leon Danon. “Community Structure in Jazz.” In: *Advances in Complex Systems* 6 (2003), p. 565.

[7] Alain Gély et al. “Uncovering and Reducing Hidden Combinatorics in Guigues-Duquenne Bases.” In: *ICFCA*. Ed. by Bernhard Ganter and Robert Godin. Vol. 3403. Lecture Notes in Computer Science. Springer, 2005, pp. 235–248.

[8] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. “Signed networks in social media.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2010, pp. 1361–1370.

[9] M. Lichman. *UCI Machine Learning Repository.* 2013.

[10] David Lusseau et al. “The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations.” In: *Behavioral Ecology and Sociobiology* 54.4 (2003), pp. 396–405.

[11] Juraj Macko. “On Open Problem — Semantics of the Clone Items.” In: *Contributions to the 10th International Conference on Formal Concept Analysis (ICFCA 2012).* Ed. by Florent Domenach, Dmitry I. Ignatov, and Jonas Poelmans. Vol. 876. CEUR Workshop Proceedings. CEUR-WS.org, 2012, pp. 130–144.

[12] O. L. Mangasarian and W. H. Wolberg. “Cancer diagnosis via linear programming.” In: *SIAM News* 23.5 (1990), pp. 1–18.

[13] Raoul Medina, Lhouari Nourine, and Olivier Raynaud. “Interactive Association Rules Discovery.” In: *ICFCA*. Ed. by Rokia Missaoui and Jürg Schmid. Vol. 3874. Lecture Notes in Computer Science. Springer, Jan. 27, 2006, pp. 177–190.

[14] Raoul Medina et al. “Efficient algorithms for clone items detection.” In: *In CLA’05*. 2005, pp. 70–81.

[15] Tore Opsahl and Pietro Panzarasa. “Clustering in weighted networks.” In: *Social networks* 31.2 (2009), pp. 155–163.

[16] Ryan A. Rossi and Nesreen K. Ahmed. “The Network Data Repository with Interactive Graph Analytics and Visualization.” In: *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*. 2015.

[17] Stanley Wasserman and Katherine Faust. *Social Network Analysis. Methods and Applications.* Structural Analysis in the Social Sciences. New York, USA: Cambridge University Press, 1994.

[18] Rudolf Wille. “Ordered Sets: Proceedings of the NATO Advanced Study Institute held at Banff, Canada, August 28 to September 12, 1981.” In: ed. by Ivan Rival. Dordrecht: Springer, 1982. Chap. Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts, pp. 445–470.