

Ekvivalence a blokové modelování v analýze sociálních sítí^{1,2}

Mgr. Tomáš Diviák, Filozofická fakulta, Univerzita Hradec Králové; Filozofická fakulta, Univerzita Karlova a ICS při University of Groningen

Equivalence and Blockmodeling in the Analysis of Social Networks

Extended abstract:

The paper presents to Czech social scientists an introductory review of the concept of equivalence and the method of blockmodeling in social network analysis (SNA). After introducing the central concepts of SNA such as node and tie, along with their basic metrics such as centrality and cohesion, I present the concepts of role and position. These are treated by SNA as clusters of nodes with similar ties, something I juxtapose to algorithms to identify cohesive subgroups of nodes. Subsequently, I define and compare the two most frequently applied types of equivalence – structural, which is strict but broadly applicable, and regular, which is more liberal but has limited uses. Structural equivalence builds on a strict definition of similarity of ties, treating as equivalent only such nodes that have the same ties to the same other nodes. Regular equivalence works with looser criteria and better corresponds with both the theoretical and the intuitive notions of role; this, however, is outweighed by the absence of a unique regular-equivalent solution within a network and by the difficulty to process networks with undirected ties. Regular-equivalent nodes are such that have ties to other mutually equivalent nodes. I present examples to demonstrate the differences between both definitions. In the following section, I discuss measurement of similarity between the different nodes' profiles of ties (e.g., correlation and Euclidean distance) and possible uses of the standard statistical methods of cluster analysis and multidimensional scaling to detect equivalent classes of nodes within networks. After pointing to the weaknesses of these techniques in network data analysis, I present blockmodeling as a method designed specifically to identify roles and positions within networks. I schematize the blockmodeling procedure and present its basic terms before comparing classic inductive blockmodeling, which is primarily fit for the purposes of exploration and network reduction, with deductive generalized blockmodeling, which is applicable in testing hypotheses about basic structural characteristics of a network. I bring attention to the strengths and weaknesses of both approaches. Relatedly, I present an application of blockmodeling especially for the purposes of simplified network representation, comparing structural patterns across networks, and testing structural theories. In the following section, I demonstrate specific blockmodeling algorithms based on both structural equivalence (CONCOR and Tabu Search optimization) and regular equivalence (REGES and Tabu Search optimization). Then I verify the adequacy of their resulting assignment of positions to nodes using eta coefficient, Q modularity and correlation of the ideal blocked and the empirical adjacency matrices. In the concluding section, I demonstrate the entire blockmodeling procedure on an empirical case of a small network with undirected ties using the UCINET software tool, including interpretation of results. Finally, I reflect the contemporary position of blockmodeling among leading research approaches in SNA, referring to other empirically oriented studies that demonstrate the broad applicability and utility of position analysis.

Keywords: social network analysis, blockmodeling, cluster analysis, relational data

<http://dx.doi.org/10.13060/1214438X.2017.1.15.366>

Ačkoliv je za našimi hranicemi analýza sociálních sítí (dále SNA³) dobře etablovaným okruhem metod či dokonce subdisciplínou v rámci sociálních věd s prestižními konferencemi, vysoce profilovanými odbornými periodiky a výzkumnými centry na renomovaných univerzitách, v českých luzích a hájích se jedná spíše o příslovečnou popelku. Síťovou perspektivu ve svých studiích aplikovalo doposud nemnoho autorů, v posledních letech se však situace v tomto ohledu zlepšuje – namátkou lze uvést práce Vašáta s Bernardem [2015] o egocentrických sítích imigrantů a Císaře a Navrátila [2014] o sítích sociálních hnutí. Nelze se ubránit dojmu, že je to svým způsobem škoda vzhledem k silným základům v Petruskově [1969] průkopnické monografii o sociometrii (de facto

předchůdkyni SNA). Ambicí této statě je přispět do českého prostředí úvodním didaktickým textem o konceptu a metodě, která jde nad rámec prostého popisu sítí pomocí měr koheze či měr centrality, a přiblížit tak českým sociálním vědcům a vědkyním pokročilejší, byť stále relativně snadno přístupné partie SNA.

Základní pojmy SNA

Zde jen v krátkosti shrnu základní pojmy, s nimiž pracuje SNA a které pak dále užívám v tomto článku a jejichž znalost je tak pro pochopení tohoto textu nezbytná. Jako podrobnější úvodní literaturu doporučuji klasický text Wassermana

a Faustové [1994], online dostupnou učebnicí Hannemana a Riddla [2005] nebo z novějších publikací zdařilou knihu Borgattiho, Everetta a Johnsona [2013]. I v češtině jsou k dispozici poněkud skromnější, avšak stále názorné a dostatečně detailní práce poskytující úvod do SNA, např. kapitola Šubta se Schmidem [2010] v edici Soudobá sociologie nebo na aplikaci SNA v policejním vyšetřování zaměřené uvedení do problematiky od Mazáka, Homolové, Diviáka a kolegů [2015].

Naprostý základ SNA představují pojmy uzlu a vazba. Uzly představují libovolné entity a v grafu jsou reprezentovány body. Jako uzly mohou figurovat jedinci, ale třeba i skupiny, organizace či státy, přičemž pro jednoduchost v jedné síti připouštíme pouze uzly jednoho druhu (tj. pouze státy, pouze organizace či pouze jedince)⁴. Vazby pak spojují uzly mezi sebou, v grafu jsou reprezentovány úsečkami. Pro vazby jsou určující dvě vlastnosti – *orientace a síla*. Orientované vazby jsou takové, u nichž lze popsat, od kterého uzlu ke kterému směřují (např. jeden jedinec jinému něco dává, jeden stát posílá druhému peníze apod.). O vazby neorientované se jedná v opačném případě (např. členství ve stejné politické straně je vazbou neorientovanou). Síla vazeb zase značí těsnost či intenzitu daného vztahu. Vazby lze takto rozlišit na binární (indikují pouze přítomnost/absenci vazby), ordinální (míra souhlasu jednoho člověka s postoji druhého), kardinální (třeba množství předaných peněz) nebo populárně na silné a slabé [Grannovetter 1973]. Ve finále tak lze mít vazby orientované binární, neorientované kardinální atp. Uzly spojené vazbami dohromady tvoří síť.

Míry centrality jsou asi nejznámějším a nejpobulárnějším konceptem v SNA [Morselli 2009: 38]. Jedná se o sadu metod sloužících k určení relativního významu daného uzlu v síti. Přestože se všechny míry centrality vztahují k jednomu konceptu (významu uzlu v síti), není ani zdaleka nadbytečné počítat a porovnávat jich více najednou, protože se každá z měr centrality tohoto konceptu dotýká z jiného úhlu a vzájemně jsou tedy komplementární. Nejzákladnější z měr centrality je *stupeň (degree)*⁵, který udává prostý součet všech vazeb (a v případě ohodnocených vazeb i jejich hodnot), daného uzlu. V případě sítí s orientovanými vazbami lze rozlišit dvě složky stupně – *příchozí stupeň (indegree; počet příchozích vazeb)* a *odchozí stupeň (outdegree; počet vazeb odchozích)*. *Mezilehlost (betweenness)* uzlu oproti tomu vyjadřuje poměr nejkratších cest mezi všemi možnými dvojicemi ostatních uzlů, na nichž daný uzel leží, ke všem ostatním. Je tedy numerickým vyjádřením přemostování a spojování – uzel s vysokou hodnotou mezilehlosti je jakýmsi gatekeeperem, přes nějž proudí množství komodit v dané síti (ať už to jsou peníze, informace, nebo cokoli jiného). Tyto dvě nejběžnější míry centrality lze ještě doplnit celou řadou dalších, jako jsou např. blízkost (closeness) nebo eigenvector. Jejich výklad však pro následující text není páteří a není zde na něj ani dostatek místa, pročez zájemce odkazují na výše jmenované publikace, kde je možné dozvědět se více.

K popisu sítí jako celku slouží *míry koheze*. Patrně nejpoužívanější a pro další výklad v tomto textu nejzásadnější z nich je *hustota*. Hustota vyjadřuje poměr vazeb, které v síti reálně jsou navázány, ke všem vazbám, které by v této síti bylo možné navázat. Nabývá hodnot od nuly (síť bez jakýchkoliv vazeb) do jedné (síť, kde má každý uzel vazbu ke všem ostatním). Konceptem na pomezí měr centrality a měr koheze je *centralizace*

sítě. Centralizace vyjadřuje, do jaké míry se centralita (např. stupeň) soustředí v nejcetrálnějších uzlech v síti. Maximálně centralizovaná síť vypadá jako hvězda, v jejímž středu je jeden uzel, který má vazbu na všechny ostatní uzly a žádné další vazby se v ní již nevyskytují. Centralizace pak vlastně říká, nakolik se pozorovaná síť podobá síti hvězdy. Konečně, *geodetická vzdálenost* udává nejmenší počet vazeb, který leží mezi danou dvojicí uzlů. Jako míra koheze slouží průměrná geodetická vzdálenost – čím je menší, tím je síť kompaktnější.

Posledním okruhem základních technik v SNA je *identifikace podskupin*⁶ v síti, tj. uskupení těsně spojených uzlů. Základním typem takové podskupiny je *klika*. Definována formálně, klika je maximální kompletní podgraf. De facto to znamená, že klika je takovou skupinou uzlů (minimálně tříčlennou), v níž musí mít každý uzel vazbu ke všem ostatním uzlům z této skupiny, a tudíž je uvnitř kliky maximální hustota vazeb. Maximální znamená, že do této podskupiny nelze přidat žádný další uzel, aniž by došlo k porušení pravidla kompletnosti, tedy přítomnosti všech možných vazeb.

Všechny analýzy zde provedené i algoritmy zde popsané odpovídají jejich implementaci a užití v programu pro SNA zvaném UCINET [Borgatti et al. 2002]. Tento program využívám z několika důvodů – je intuitivní k používání (nevyžaduje tak znalosti programování jako knihovny pro SNA v R), má velice dobře zpracovaný manuál a obsahuje všechny relevantní metody. Pro hlubší zájemce lze vyjma R doporučit ještě *Pajek*, který není uživatelsky tak přátelský jako UCINET a nemá tak široký záběr, pro blokové modelování nicméně nabízí velké množství algoritmů a procedur⁷.

Ekvivalence a její druhy

Pojmy status a role jsou úhelnými kameny sociologické teorie. V sociologii, ale i příbuzných oborech, patří tyto termíny mezi nejdiskutovanější již od poloviny minulého století [Borgatti et al. 2013: 206]. V síťové perspektivě je tato problematika zakotvena prostřednictvím konceptu ekvivalence. Ačkoliv lze rozlišit několik druhů ekvivalence s vlastní definicí a způsoby měření, nezanedbatelnou výhodou SNA spatřuji v tom, že jsou tyto definice přesné a jednoznačné. Je nicméně třeba mít na paměti, že tato přesnost a jednoznačnost se váže k poměrně úzkému pojetí v rámci sítí a nemůže tak suplovat řadu tradičtějších pojetí a definic, které v sociálních vědách pojmy status a role mají.

V SNA je obvykle užíván termín pozice jako synonymum pro termín status. *Pozice* je v SNA označení pro „souhrn individuů, která jsou podobně zakotvena v síti vztahů“ [ibid.]. *Role* jsou naproti tomu síťově chápány jako „vzorce vztahů, které se uplatňují mezi aktéry nebo mezi pozicemi“ [ibid.]. Na první pohled se může zdát, že jsou pozice jen dalším způsobem, jak identifikovat podskupiny. To by ovšem nebylo přesné, neboť na rozdíl od hledání shluků vzájemně propojených uzlů (tj. identifikace kohezivních podskupin) poziční přístup popisuje takové třídy⁸ uzlů, které mají stejné či podobné strukturální vlastnosti, tedy podobné vztahy k jiným uzlům a jejich třídám nebo i mezi sebou navzájem [Borgatti et al. 2013: 207]. Uzly, jež jsou si ekvivalentní, přitom ale k sobě vůbec žádné vazby mít nemusejí, což ale právě bývá definičním znakem podskupin. Nejlepší bude ilustrovat tento rozdíl na příkladu. Třeba základní školu si lze jednoduše představit jako síť, v níž jednou

z pozic je pozice učitele. Aktéři v tomto postavení se na jedné straně vyznačují tím, že mají vztah k pozici žáků (učí je, kárají je atp.) a na straně druhé se zodpovídají řediteli, aniž by se však mezi sebou navzájem nutně zodpovídali jeden druhému, kárali se nebo se učili jeden od druhého⁹.

Proč se tedy zabývat studiem ekvivalence v sítích? Borgatti a jeho kolegové [2013: 207–208] elegantně shrnují tři hlavní důvody. Jednak řada výzkumů poukazuje na to, že strukturně ekvivalentní aktéři vykazují vyjma strukturních podobností také podobnosti v postojích či chování – příslušnost ke stejné pozici tak může indikovat homogenitu i v jiných nerelačních atributech aktérů. Jinými slovy, s ekvivalencí se úzce pojí mechanismus homofilie. Dalším důvodem je to, že ekvivalentní uzly jsou záměnné – jejich záměnou se struktura sítě nezmění. Případně odebrání libovolného uzlu ze sítě může kompenzovat jemu ekvivalentní partner, což má implikace pro všechny, kteří chtějí sítě (de)stabilizovat (např. ve výzkumu organizací či při policejním vyšetřování). Konečně, zachycením celé sítě prostřednictvím malého počtu pozic a rolí mezi nimi lze efektivně redukovat síť, aniž by přitom došlo ke ztrátě jejích podstatných strukturních charakteristik, což je užitečné třeba při analýze sítí s velkým počtem uzlů.

Rozlišuje se mezi několika druhy ekvivalence, nejzásadnější a nejvíce používané jsou však dva, a to strukturní a regulární¹⁰. Prvním a dá se říci výchozím typem je *ekvivalence strukturní*. Byla definována a popsána Lorrainem a Whitem [1971]. Obecně lze říci, že za strukturně ekvivalentní považujeme takové dva uzly, které mají *vazby o stejné intenzitě a ke stejným uzlům* [Hanneman, Riddle 2005]. Jak jsem již poukázal výše, uzly, které jsou strukturně ekvivalentní, jsou z hlediska struktury sítě nerozlišné, a tudíž vzájemně zaměnitelné [Borgatti et al. 2013: 208]. Záměnou uzlů ze stejné ekvivalentní třídy se tedy síť nijak nezmění. Toto pojetí ekvivalence je velmi „silné“ a restriktivní, neboť na to, aby byly dva uzly strukturně ekvivalentní, je třeba splnit dost přísné požadavky. V praxi však k takto přesné podobnosti ve struktuře a síle vazeb dochází spíše vzácně, a proto byla vytvořena řada metod a algoritmů vyjadřujících míru jejich podobnosti [Schmidt, Šubrt 2010]. Nejčastěji bývají tyto metody založeny na porovnávání podobnosti profilů jednotlivých uzlů, tedy na porovnávání podobnosti řádků a sloupců v datové matici. Následně se tyto (ne)podobnosti zpracovávají pomocí klasických statistických metod pro klasifikaci objektů (typicky klastrová analýza) nebo na jejich principech založených zvláštních procedurách, o nichž se rozepisují níže [Borgatti et al. 2013: 210–223; Wasserman, Faust 1994: 363–397]. Příslušnost ke stejnému shluku či podobné umístění v grafickém znázornění (např. u multidimenzionálního škálování) pak ukazují na příslušnost ke stejné pozici.

Wasserman a Faustová [1994: 468–469] však upozorňují i na nevýhody a omezení, která s sebou nese uchopení ekvivalence jejím strukturním způsobem. Omezení pouze na identické vazby a aktéry neumožňuje jednak obecnou formalizaci teoretického konceptu sociální pozice a jednak neumožňuje porovnávat role a pozice napříč populacemi. Argumentují tím, že myšlenka role je obecným konceptem, kterýžto je nezávislý na identitě konkrétních zahrnutých individuí, a je proto nutné mít možnost popsat a porovnat zobecněné a abstraktní vlastnosti vazeb aktérů, aniž by přitom bylo nezbytné omezovat se právě jen na jednotlivé aktéry a jejich vazby. Pro ilustraci se zde vrátím k výše uvedenému příkladu základní školy.

Omezená definice strukturní ekvivalence např. neumožňuje, aby byly ve stejné pozici dva učitelé, kteří učí sousední třídy, protože mají vazbu (tj. vyučují) k jiným konkrétním aktérům, třebaže jsou všichni tito aktéři žáci a jsou vyučováni stejným způsobem. Z hlediska intuice i sociologické teorie je ale právě možnost zahrnout oba učitele do stejné pozice významná, a to nehledě na skutečnost, zda se jedná o učitele jménem Novák, nebo učitele jménem Dvořák – oba jsou koneckonců učiteli, což je jejich společný a podstatný znak.

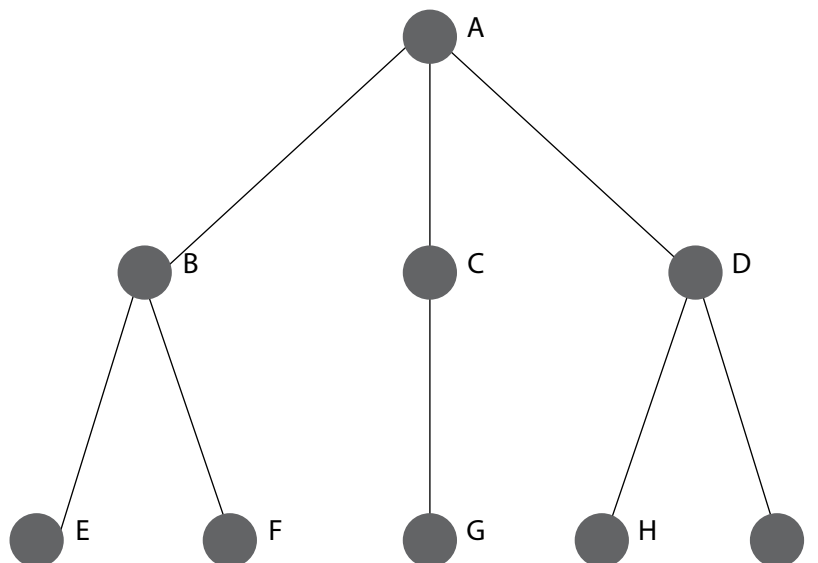
Slabším a výrazně méně restriktivním typem ekvivalence je *ekvivalence regulární* [White, Reitz 1983]. Tento druh nevyžaduje identické vazby k identickým aktérům ani zachování vzdáleností a struktury v síti. Stačí, že uzly ve stejné pozici mají stejné vazby k jiným uzlům, které jsou zároveň také ve stejné pozici [Wasserman, Faust 1994: 473]. Jakkoliv se tato obměna jeví jako kosmetická, je ve skutečnosti zcela zásadní. Aktéři v regulárně ekvivalentních pozicích totiž nemusí spadat do pozic či lokací v síti s ohledem na jiné individuální aktéry, ale mají spíše stejný typ vztahů k některým členům jiných pozic [Hanneman, Riddle 2005]. Regulární ekvivalence si tak ze strany sociologů zaslouží zvláštní pozornost, neboť ve své volnosti nejlépe zachycuje ideu sociální role institucionalizované prostřednictvím normami a sankcemi vymezených vztahů k nositelům jiných rolí, což je jádrem sociologické perspektivy [ibid.]. Na regulární ekvivalenci je zvláštní také fakt, že se v síti může vyskytovat více možných způsobů, jak uzly v ní rozdělit do ekvivalentních tříd, přičemž se rozdělení s nejmenším počtem tříd označuje jako *maximální regulární ekvivalence* [Wasserman, Faust 1994: 475]. Toto pojetí ekvivalence má sice širokou uplatnitelnost, protože jej lze aplikovat na jedno- a vícerelační sítě, vazby ohodnocené a binární, ale algoritmy k jeho detekci mají problém se sítěmi s neorientovanými vazbami bez izolátů (tj. uzlů bez vazeb), na kterých jako maximální regulární ekvivalenci detekují celou síť, což však lze jen stěží označit za závěr jakkoliv zajímavý nebo přínosný [Wasserman, Faust 1994: 475; Everton 2012: 294].

Výhodou regulární ekvivalence oproti strukturní je to, že strukturně ekvivalentní uzly kvůli tomu, že mají stejné vazby ke stejným aktérům, spadají také pravděpodobně do stejných kohezivních podskupin, což ale pro regulárně ekvivalentní uzly nutně neplatí, a tak lze tímto způsobem lépe zachytit role a pozice nezávisle na (pod)skupinové příslušnosti [Borgatti et al. 2013: 223]. Vypomohu-li si opět výše použitým příkladem, pak z hlediska regulární ekvivalence jsou si učitelé Novák a Dvořák ekvivalentní bez ohledu na to, kolik žáků a jak vyučují, protože důležité je pouze to, že někoho vyučují, tj. mají podobný typ vazeb k uzlům, které jsou si též ekvivalentní (žáci). Nevýhodou je především nejednoznačnost ve formě většího počtu řešení a již zmíněné obtíže s rozřazováním neorientovaných sítí.

Graf č. 1 je dnes již klasickým didaktickým příkladem pro demonstraci odlišností mezi různými způsoby, jak definovat ekvivalenci. Na obrázku je celkem sedm strukturně ekvivalentních pozic. Žádný uzel nemá stejné vazby jako uzel A, a ten je tak pozicí sám o sobě. Uzly B, C, D sice mají všichni vazbu k A, nicméně jejich další vazby jsou bez jakéhokoliv překryvu, a proto i každý z těchto uzlů samostatně tvoří pozici. Uzly E a F mají každý jen jedinou vazbu – na uzel B. Žádnou další vazbu k jiným uzlům nemají – tyto dva uzly jsou tedy strukturně ekvivalentní. Stejně je tomu s uzly H a I. Poslední neza-

Graf 1.: Ilustrace ekvivalentních tříd [Wasserman, Faust 1994: 468]

Figure 1: Illustration of equivalent classes [Wasserman, Faust 1994: 468]



řazený uzel je G, který nemá žádný strukturně ekvivalentní protějšek, a je tak opět pozicí sám o sobě. Z hlediska regulární ekvivalence jsou v síti na obrázku č. 1 přítomny tři ekvivalentní třídy. Nejlépe je to patrné vizuálně od spodu grafu. První pozice se skládá z uzlů E, F, G, H a I. Regulárně ekvivalentní jsou jednak proto, že samy mezi sebou vazby nemají, a pak také proto, že každý z nich má alespoň nějakou vazbu na trojici B, C, D (což je regulárně ekvivalentní pozice) a žádný z nich nemá vazbu k uzlu A. Uzly B, C a D jsou regulárně ekvivalentní díky svým vazbám ke spodní pozici, absenci vzájemných vazeb a vazbě na uzel A, který samostatně tvoří poslední, třetí pozici. Jak je vidno, vymezení pozic probíhá v rámci regulární ekvivalence současně. Je rovněž patrné, že regulární ekvivalence konvenuje více intuitivní představě o ekvivalenci, zatímco strukturní ekvivalence představuje přísněji definované pojetí tohoto konceptu. Mimochodem, každé strukturně ekvivalentní řešení je zároveň také regulárně ekvivalentní (nikoliv však maximálně regulárně; [Everton 2012: 294]).

Měření podobnosti

Jak jsem již předeslal výše, perfektní ekvivalence jakéhokoliv typu se v praxi téměř nevyskytuje, a je tedy potřeba měřit, jak moc jsou i uzly ekvivalentní, tj. do jaké míry se jejich vazby podobají. V případě ekvivalence strukturní k tomu lze užít standardní míry porovnávací podobnost dvou vektorů (řádků či sloupců v datové matici), jakými jsou např. *korelace* nebo *eukleidovská vzdálenost* [Borgatti et al. 2013: 211], popř. poměr přesně se shodujících vazeb pro binární vazby a míry ordinální asociace pro ordinálně hodnocenou sílu vazeb v dané síti [Wasserman, Faust 1994: 475]. Korelační koeficienty či eukleidovské vzdálenosti lze pak pojmout jako vstupní data pro *multidimenzionální škálování* nebo různé procedury *klastrové analýzy* [ibid.: 214]. Náleží-li uzly do stejného klastru nebo jsou-li si blízko v mapě multidimenzionálního škálování, pak je lze označit za ekvivalentní. Je nicméně nutné přihlídnout k faktu, že míry podobnosti nemusejí nutně mít stejné výsledky (tj. korelace = 1 a eukleidovská vzdálenost =

0), a je tak dobré se rozhodnout, kterou z nich v analýze zvolit. Eukleidovské vzdálenosti reflektují méně strukturní ekvivalence, když se liší uzly v průměrech a rozptylech svých vazeb [Wasserman, Faust 1994: 374]. Při standardizaci dat se problém eukleidovských vzdáleností odstraní a procedury fungují jako korelace na nestandardizovaných datech [ibid.], což mne však vede k názoru, že je obecně lepší užívat korelace. Výjimkou je případ tzv. *multiplexních sítí*¹¹, v nichž nelze pomocí korelací podobnosti zachytit, zatímco výpočet eukleidovských vzdáleností to nikterak neovlivňuje.

V případě regulární ekvivalence se podobnosti mezi uzly měří poněkud odlišně – prostřednictvím konceptu *sousedství v síti* (u sítí s orientovanými vazbami) a *geodetických vzdáleností* (u sítí s neorientovanými vazbami; [Hanneman, Riddle 2005]). Sousedstvím uzlu v grafu jsou všechny uzly, které mají k danému uzlu vazbu [Wasserman, Faust 1994: 476]. Uzly z jedné regulárně ekvivalentní pozice lze takto definovat jako takové, které mají ve svém sousedství uzly (alespoň však jeden) z jiné regulárně ekvivalentní pozice [ibid.]. Rozhodnutí, zda jsou dva aktéři regulárně ekvivalentní, vyžaduje zároveň posouzení, zda jsou i jejich sousedi regulárně ekvivalentní. V případě čisté regulární ekvivalence to znamená, že jsou-li jakékoliv dva uzly *i* a *j* v ekvivalentní pozici, vazba uzlu *i* k libovolnému uzlu *k* z jiné pozice implikuje, že uzel *j* musí mít zároveň vazbu k uzlu *k*, kterýžto náleží do stejné regulárně ekvivalentní pozice jako uzel *k*. Pro všechny regulárně ekvivalentní dvojice uzlů tudíž platí, že se jejich vazby k uzlům z jiných pozic musejí shodovat tímto způsobem, tj. má-li jeden uzel vazbu k jinému uzlu z jiné pozice, pak má vazbu na kterýkoliv uzel z této pozice i jemu regulárně ekvivalentní uzel [ibid.: 479]. V případě empiricky „nečisté“ regulární ekvivalence pouze algoritmus [podrobněji vysvětlený v Borgatti, Everett, Freeman 2002] porovnává podobnost dvojic uzlů pomocí podobnosti ve skladbě jejich sousedství či v geodetických vzdálenostech k ostatním uzlům (tj. čím více shodných sousedů dva uzly mají nebo čím podobněji jsou vzdáleny k ostatním, tím jsou si podobnější). Z těchto podobností následně vzejde symetrická matice a ta se pak jednoduše může stát vstupem pro klastro-

vou analýzu nebo multidimenzionální škálování podobně jako v případě ekvivalence strukturní.

Ekvivalentní třídy lze tedy odhalit i za pomoci dnes již běžně používaných statistických metod. Výhodou především klastrové analýzy je její všeobecná známost, a tedy fakt, že ji umějí aplikovat i výzkumníci, kteří ještě nemuseli přijít do kontaktu se specifickými metodami SNA. V tomto ohledu se jeví nejlepší klastrování metodou *complete link*, která není náchylná na řetězení objektů do shluků a vytváří homogennější a stabilnější klastry [Wasserman, Faust 1994: 381]. Klastrová analýza má ovšem i své nevýhody při analýze běžných atribučních dat, které se vztahují i na její užití na data relační. Jako nejzávažnější se jeví zejména odhalení klastrů na základě zanedbatelných rozdílů mezi objekty, takže dochází k objevení shluků i tam, kde reálně vlastně vůbec nemusejí být nebo kde se nachází jeden jediný homogenní shluk [Hendl 2009: 496]. Specifickou nevýhodou užití standardní hierarchické klastrové analýzy v případě analýzy sítí je komplikace u orientovaných sítí, kdy běžně užívané statistické programy berou v potaz pouze řádky v matici nepodobností, tedy pouze odchozí vazby (odchozí stupeň), proto je nutné pracovat buď s maticí sousednosti, nebo vektor (ne)podobností prodloužit [Robins 2015: 187]. To však není vždy snadno (či dokonce vůbec!) proveditelné, zatímco softwarové balíky pro analýzu sítí tento problém řeší zpravidla automaticky. Je zde ještě další drobná nevýhoda, kterou však lze s trochou pile ze strany výzkumníka odstranit – klastrová analýza sice dokáže rozlišit uzly do ekvivalentních pozic, nedokáže však určit, jaké role (tj. vazby) mají uzly v nich k jiným uvnitř jedné pozice nebo napříč pozicemi. To však lze automaticky určit pomocí blokového modelování.¹²

Blokové modelování

Podle Wassermana a Faustové [1994: 363] by měl postup při analýze rolí a pozic v síti vypadat zhruba takto: 1) formální definice ekvivalence, 2) změření, do jaké míry uzly a vazby splňují tuto definici, 3) reprezentace ekvivalentních tříd a 4) stanovení adekvátnosti reprezentace. V prvním kroku je nutné zvolit teoreticky vhodný typ ekvivalence. Následně se zvolí odpovídající metoda měřící zvolený typ ekvivalence v síti. Reprezentace ekvivalentních tříd se skládá z rozkladu sítě na třídy (podmnožiny uzlů) a popisu vztahů mezi nimi a uvnitř nich (jakou vazbu má která pozice ke které). Konečně, výslednou zjednodušenou reprezentaci pozic a rolí je nutné porovnat s původními daty a stanovit tak, do jaké míry na výchozí data „sedí“ a zda je tedy hodnověrně reprodukuje. První krok jsem již popsal výše v části o ekvivalenci a jejích typech. Druhý krok jsem již také popsal, a to v části o měření podobnosti mezi uzly. Stěžejní třetí krok vztahující se k reprezentaci ekvivalentních tříd popíši nyní prostřednictvím představení konceptu modelování bloků. K ověření adekvátnosti reprezentace se vyjadřuji posléze.

Změřením (ne)podobností mezi uzly a následným rozdělením uzlů do pozic (např. pomocí klastrové analýzy) lze přerovnat matici dat tak, aby v ní byly uzly seřazeny vedle sebe podle příslušnosti k jednotlivým pozicím. Již při prostém pohledu na takto přerovnanou matici začíná být mnohdy patrné, kde se v matici vyskytují vazby mezi pozicemi a kde naopak nikoliv. Na tomto základě je možné matici zjednodušit na

tzv. *image matici*, v níž jsou pouze pozice a vztahy mezi nimi nebo uvnitř nich. Takovouto maticí není pak již nijak složité vyjádřit i v grafické podobě jako tzv. *image graf*. Proces, jehož prostřednictvím vzniká z původní matice dat její zjednodušená reprezentace, se nazývá blokové modelování [Borgatti et al. 2013: 216]. Blokem se rozumí dvě pozice a všechny vazby mezi aktéry k nim náležející [Batagelj et al. 2011]. *Blokmodel* zahrnuje jednak rozklad aktérů v síti na diskrétní podmnožiny (tedy pozice) a jednak stanovení přítomnosti či absence vazby mezi každou dvojicí pozic či uvnitř pozice [Wasserman, Faust 1994: 395]. V síti s počtem bloků rovnajícím se k existuje potom k^2 bloků, pomocí nichž lze zachytit strukturu celé sítě ve zjednodušeném blokmodelu [Batagelj et al. 2011]. Bloky jsou dvojího druhu – buď 1-bloky, které zahrnují existující vazby (v binárních sítích tedy samé 1 – odtud jejich název), nebo 0-bloky, analogicky nezahrnující žádné vazby (samé 0 v binárních sítích; [Borgatti et al. 2013: 216]). Užitečnost blokmodelů tkví především v jejich možnosti teoretické interpretace, validace strukturálních teorií a porovnávání strukturních vzorců napříč sítěmi [Wasserman, Faust 1994: 394].

Obecně se rozlišují dva přístupy k blokovému modelování – klasické a generalizované [Batagelj et al. 2011]. Při klasickém blokovém modelování se nejprve provádí měření (ne)podobností, na jejichž základě jsou uzly pomocí různých algoritmů rozřazeny do pozic, protože bývá tento přístup také označován jako nepřímé blokové modelování [ibid.]. Generalizované (přímé) blokové modelování je oproti tomu založené na konstrukci ideálního blokmodelu a jeho následném porovnání s empirickým, kdy se poměruje míra jejich odlišnosti (detailně popsáno v [Doreian et al. 2005]). Ačkoliv se tvůrci generalizovaného modelování bloků snažili napravit nedostatky klasického blokmodelování, především jeho omezení na induktivní užití a arbitrárnost klastrovacích procedur [Batagelj et al. 2011], a navrhnout optimální postup pro malé sítě, jejich přístup je aplikovatelný pouze na binární vazby, což je značně limitující. Žiberna [2007; 2008; 2009] se sice systematicky pokusil tato omezení překonat v intencích jak strukturní, tak regulární ekvivalence, avšak s prozatím ne zcela uspokojivými výsledky, neboť jím navržené postupy by vedly k ne vždy optimálním výsledným blokům [blíže v Batagelj et al. 2011]. Jsem navíc přesvědčen, že významnější nedostatek, který klasickému blokovému modelování vytýkají Batagelj a jeho kolegové, tedy induktivní přístup a s ním související arbitrární interpretace výsledků, je do určité míry překonatelný – jak Wasserman a Faustová [1994: 383] podotýkají, dobrá znalost teorie v dané oblasti je nejlepším vodítkem k určení optimálního počtu pozic, čímž se usměrňuje arbitrárnost při volbě klasifikační procedury a následně nevhodnějšího rozdělení uzlů do pozic. Z hlediska teoreticko-vědeckého lze navíc podotknout, že pro některé výzkumné problémy je induktivní explorativní přístup vhodnější, ať už je to v případech, kdy je teorie málo propracovaná, substantivní oblast (či data) málo prozkoumaná, nebo v případech, kdy je blokové modelování pouhým mezistupněm coby prostá redukce dat.

Jak jsem zde již naznačil několikrát, empirických sítí, v nichž je prokazatelně čistá ekvivalence, je doslova jako šafránu a určení 1-bloků a 0-bloků proto nebývá zdaleka tak jednoznačné, jako v tomto ideálním případě. Důležitým rozhodnutím v blokovém modelování je proto přisouzení 1 či 0 jednotlivým blokům a stanovení toho, které bloky mají

mezi sebou či uvnitř sebe vazbu. K tomu existuje řada kritérií, jejichž užitečnost se liší případ od případu (podrobný výčet viz [Wasserman, Faust 1994: 397–407]). Jako vhodné se jeví kritérium alfa hustoty opírající se o vzácnost výskytu čistých 1-bloků a 0-bloků. Za 1-blok považuje ty bloky, jejichž hustota je vyšší než hustota celé matice (alfa), a za 0-bloky pak ty, jejichž hustota je naopak nižší než hodnota alfa. Wasserman a Faustová [ibid.] vhodnost tohoto kritéria potvrzují závěry z výzkumů, z nichž plyne, že jsou bloky robustní při změnách alfa.

Alfa kritérium (ale i jiná kritéria, např. lean fit či perfect fit) je aplikovatelné na binární vazby. Jejich paralely pro ohodnocené vazby sledují nicméně podobnou logiku. Pro tento typ vazeb lze v image matici mít buď dichotomické, nebo i spojité hodnoty, nicméně snazší a lépe interpretovatelné jsou hodnoty dichotomické. Za zmínku stojí dvě kritéria. Prvním z nich je kritérium maximální hodnoty – za 0-blok je považován každý takový, ve kterém jsou vždy všude nízké hodnoty, zatímco kdykoliv je v bloku aspoň jedna vazba s vysokou hodnotou, je označen jako 1-blok. Určitou analogií ke kritériu alfa hustoty je u ohodnocených vazeb kritérium průměru. Podle tohoto kritéria může být za 1-blok považován takový, v němž je průměrná hodnota vazby větší či rovna průměrné hodnotě vazby v celé datové matici. Při nesplnění této podmínky se jedná o 0-blok. Pozor je jen třeba dávat na vazby o ordinální síle na škále velmi negativní – velmi pozitivní, kde může průměr zkreslovat, a je tak lepším kritériem maximální hodnota [Wasserman, Faust 1994: 407]. Existují ovšem případy zvláště menších sítí, v nichž stačí pohled na matici permutovanou do bloků a je zřejmé, které bloky jsou 1-bloky a které 0-bloky.

Metody blokového modelování a detekce pozic a míry jejich adekvátnosti

Když je zvolen vhodný typ ekvivalence, přichází na řadu volba vhodné metody či algoritmu, které provedou všechny výpočetní operace, což nakonec vyústí v blokmodel. Program UCINET nabízí hned několik způsobů, jak identifikovat role a pozice v síti. K analýze jsem jich užil víc a jejich výsledky porovnal. Zde popisují metody CONCOR, Tabu search optimizaci, aplikace klastrové analýzy a REGE.

Procedura CONCOR¹³ je, jak připouští Hanneman s Riddlem [2005], poněkud zastaralá, stále však produkuje smysluplné výsledky. Zjednodušeně řečeno, CONCOR nejprve zpracuje korelace profilů uzlů, následkem čehož vznikne korelační matice, v níž jsou vzájemně korelovány řádky znovu a znovu, dokud nevznikne matice, v níž jsou jen hodnoty +1 nebo -1 [ibid.]. Na základě tohoto opakovaného korelování rozděluje vždy velký shluk (nejprve celou síť) na dvě menší vnitřně homogenní pozice. Pro badatele obeznámené s klastrovou analýzou se zde nabízí srovnání s metodou divizivního klastrování. Tato procedura však umí vytvořit jen sudé počty (funguje jako binární strom) tím, že množinu uzlů vždy dělí na dvě menší [Wasserman, Faust 1994: 380]. To pochopitelně může značně zkreslovat. I přesto, že je CONCOR široce aplikovatelný (mj. i na ohodnocené či neorientované vazby nebo multiplexní sítě), bývá mu vytýkána určitá nespolehlivost [Hanneman, Riddle 2005].

Optimizace pomocí Tabu search je mladší sestrou CONCOR. Ačkoliv je výpočetně výrazně náročnější, snaží se rozvíjet

podobnou myšlenku jako CONCOR – seskupovat do jedné pozice uzly, které jsou si co nejvíce podobné [ibid.]. Vychází předpoklad je, že dva uzly jsou si podobné v případě, že jsou jejich profily blízké průměrnému profilu bloku, v němž se nacházejí. Tabu search optimizace se tedy snaží utvářet bloky tak, aby byl rozptýl profilů uvnitř bloků co nejmenší, přičemž zároveň se snaží minimalizovat chyby, tj. vazby, které do zvoleného bloku nezapadají (přítomné v 0-blocích). Do určité míry je zde možné rozvinout analogii s klastrovou analýzou – je-li CONCOR podobný procedurám divizivní klastrové analýzy, optimizace pomocí Tabu search se naopak podobá procedurám optimačního klastrování (např. k-means nebo PAM). Teoreticky by měla tato metoda produkovat podobná rozřazení uzlů do pozic jako CONCOR, ale prakticky tomu tak zdaleka ne vždy bývá.

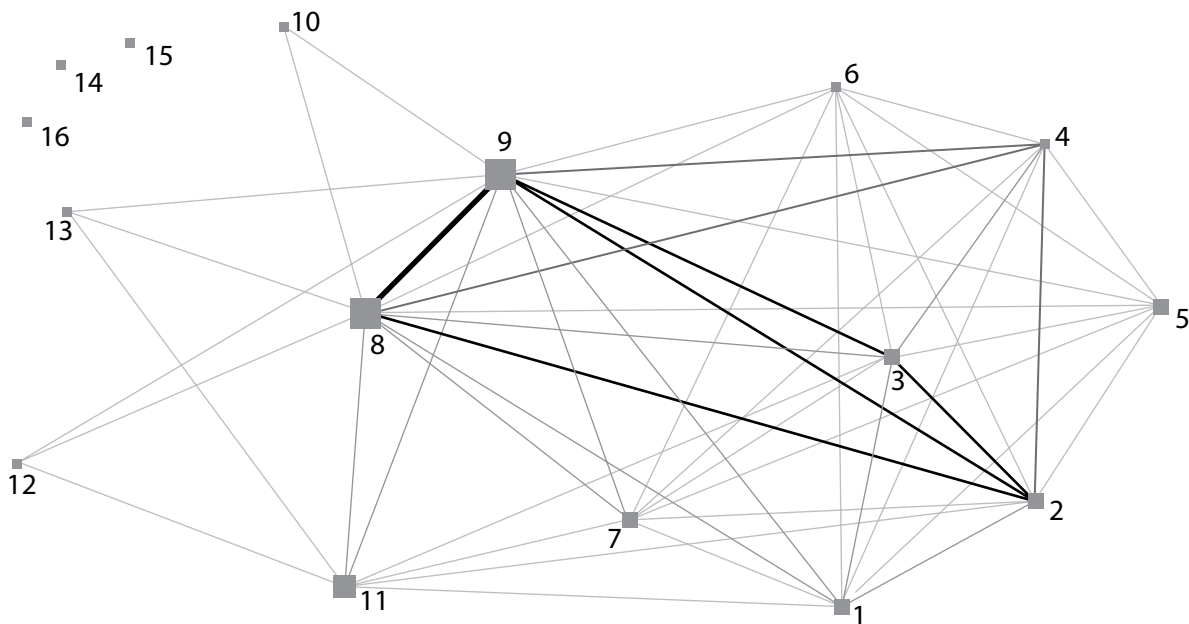
Ukázal jsem již výše, že i klastrová analýza může být využita při blokovém modelování. Nenavrhuje sice na rozdíl od CONCOR a Tabu search optimizace blokmodel ani do bloků permutovanou matici, pořad však seskupuje objekty na základě jejich podobnosti. Dalším rozdílem je také to, že obě předchozí metody nejdříve dostanou zadaný počet pozic a do něj se pak snaží co nejlépe zasadit uzly, zatímco klastrová analýza postupně spojuje podobnější uzly a je pak na badateli, kolik pozic jako řešení zvolí. Na rozdíl od CONCORu a Tabu search optimizace tedy klastrová analýza nevytváří předem zadaný počet pozic s co nejvíce podobnými uzly, ale využívá mezistupeň ve formě měření podobnosti mezi nimi a teprve poté je seskupuje do shluků. V UCINETu jsou dvě možnosti, jak provést klastrovou analýzu. První z nich je procedura *Profile Similarities*, což je vlastně klastrová analýza obohacená o to, že umí pracovat se síťovými maticemi, které umí převést na míry podobnosti (např. korelace) a následně je pomocí metody average link shlukovat [Borgatti, Everett, Freeman 2002]. Druhým způsobem, jak v programu UCINET provést klastrovou analýzu, je nejprve spočítat míry podobnosti a ty pak zadat *Johnsonovu hierarchickému klastrování* se zvolenou metodou. Doporučována bývá metoda complete link pro svou relativní odolnost vůči řetězení a větší homogenitě vytvořených klastřů [Wasserman, Faust 1994: 381].

K detekci regulární ekvivalence se užívá nejčastěji algoritmus REGE. V zásadě se jedná o extenzi principu shlukování podobnosti profilů v intencích regulární ekvivalence. Situace je nicméně komplikována tím, že určení regulárně ekvivalentní pozice vyžaduje, aby byla již určena jiná regulárně ekvivalentní pozice [Borgatti et al. 2013: 222]. To řeší REGE tak, že se iterativně přibližuje regulárně ekvivalentní matici prostřednictvím klastrování podobností sousedství (binární či multinominální sítě) nebo geodetických vzdáleností (sítě kardinální či ordinální; podrobně popsáno ve [Wasserman, Faust 1994: 479–481]). REGE je nicméně omezené svým užitím pouze na orientovaná data, a to ještě navíc s podmínkou, že se v síti musí vyskytovat alespoň jeden uzel buď s nulovým odchozím stupněm (čistý přijímač), nebo s nulovým příchozím stupněm (čistý vysílač¹⁴; [Borgatti et al. 2013: 222]).

Existuje i varianta *Tabu search optimizace* určená pro regulárně ekvivalentní blokové modelování. Vychází z definice regulární ekvivalence, když jako 1-blok chápe situaci, kdy každý uzel z jedné pozice má vazbu alespoň na jeden uzel z druhé pozice, čímž splňuje podmínku vazby regulárně ekvivalent-

Graf 2: Ilustrační síť, velikost uzlů podle hodnoty mezilehlosti a šířka vazeb podle jejich síly; pozice uzlů a vazeb manuálně mírně pozměněny za účelem lepší viditelnosti

Figure 2: Illustrative network, betweenness level indicated by size of nodes and strength of ties indicated by thickness of line; minor manual adjustments to the position of nodes and ties have been made for better visibility



ních uzlů k jiným taktéž ekvivalentním uzlům, aniž by bylo nutné, aby to byla vazba ke stejným aktérům (jak to vyžaduje strukturální ekvivalence; [Hanneman, Riddle 2005]). S tímto pravidlem pak algoritmus postupuje stejně jako v případě strukturální ekvivalence.

Posledním krokem poziční analýzy je ověření adekvátnosti získaného řešení. I k tomu existuje řada nástrojů. Využití a následně porovnat lze několik (popis podle [Borgatti et al. 2002]). První je koeficient eta, což je korelační koeficient empirické matice a ideální matice, v níž jsou jen 1 (jsou-li uzly ve stejné pozici) a 0 (v opačném případě). *Newmanova a Girvanova modularita Q* je poměr vazeb zapadajících a vazeb nezapadajících do daného řešení odečtený od tohoto poměru při náhodném rozložení vazeb. Maximální hodnota tohoto ukazatele je $1 - (1/\text{počet pozic})$. *Krackhardtův a Sternův index E-I* je rozdíl vnitřních (adekvátních) vazeb a vnějších (neadekvátních) dělený počtem všech vazeb. U *ety* i *Q* jsou žádoucí kladné hodnoty, zatímco u *EI* jsou žádoucími hodnoty záporné. U těchto procedur, u nichž je výstupem i blokmodel, je navíc k dispozici i *R2*, což může být matoucí, protože se nejedná tak úplně o klasický koeficient determinace, jak je znám z regresní analýzy, ale jedná se o korelaci získané matice permutované do bloků s její ideální obdobou, v níž jsou jen čisté 1- a 0-bloky.

Síť, která zde bude sloužit pro ilustraci výše popsaných metod, je čistě pouze demonstrační. Zahrnuje ohodnocené neorientované vazby, takže ji lze pro účely interpretace považovat třeba za síť vzájemné spolupráce mezi kolegy na libovolném pracovišti (neboť spolupráce je z definice vzájemná, také nemá směr, a tudíž ji lze v SNA zachytit jako neorientovanou vazbu). Síť je znázorněna na grafu č. 2. Již z pohledu na graf lze vyčíst, že se v ní nachází 16 uzlů, přičemž tři z nich (14, 15, 16) jsou izoláty, tedy nemají vazbu na žádný další uzel v síti. Hustota sítě činí 0,47, což znamená, že se jedná o síť poměrně dobře propojenou, neboť téměř každá druhá možná vazba je reálně navázána. Vzdálenosti v síti jsou poměrně krátké – v průměru pouze 1,359, což znamená, že od jednoho aktéra

k libovolnému druhému se lze dostat v průměru prostřednictvím necelé jedné a půl vazby. To platí samozřejmě pouze pro aktéry neizolované. Celkově lze síť tedy charakterizovat jako kompaktní. Z hlediska centrality se v síti nacházejí tři vysoce centrální aktéři 8, 9 a 11, přičemž vysoký stupeň má ještě i uzel 7. Zbylé uzly jsou tak spíše marginální. Ačkoliv si míry centrality zaslouží při běžné analýze sítí detailní pozornost, zde nejsou meritorní a pro přiblížení situace postačí takto stručný slovní popis. Vzhledem k malému počtu uzlů a neorientovaným vazbám v síti se jeví vhodnější užití strukturální ekvivalence. V obecné rovině je však následující postup použitelný i v případě blokového modelování v intencích ekvivalence regulární.

U explorativní analýzy libovolných dat bývá obecně považováno za dobrý postup vyzkoušet několik variant vhodné metody či přímo několik různých metod a jejich výsledky mezi sebou porovnat (např. [Hanneman, Riddle 2005] či [Borgatti et al. 2013: 181-206]). Na základě podobností mezi výsledky a se zřetelem na snadnou a smysluplnou interpretaci je možné vybrat finální řešení, které je užitečné vzhledem k povaze výzkumného problému. K interpretaci jsem zvolil u všech následujících metod řešení se čtyřmi klastry, resp. pozicemi. Vzhledem k tomu, že chci řešení mezi sebou porovnávat, a vzhledem k tomu, že procedura CONCOR vytváří pouze sudé počty pozic, se tento počet jeví jako rozumný ve srovnání s dvěma pozicemi, které zpravidla příliš nediferencují, a šesti pozicemi, u nichž se naopak projevují sklony k řetězení shluků a hledání rozdílů tam, kde se reálně nevyskytují (oddělování vzájemně si podobných uzlů). Z hlediska měř adekvátnosti vykazovala všechna čtyřpoziční řešení nejvyšší kvalitu. To je konzistentní i s Robinsonovým názorem [2015: 187], který na základě vlastní zkušenosti doporučuje čtyři až šest pozic.

Nejllepší je, že při porovnávání výsledků několika různých metod opřít o „pevný bod“, je-li takový k dispozici. V tomto případě vycházím z úvahy, že v síti s ohodnocenými vazbami jsou strukturálně ekvivalentní takové uzly, jež mají naprosto

Tabulka 1: Ekvivalenční třídy uzlů v síti podle procedur CONCOR, Tabu Search optimalizace, klastrové analýzy podobnosti profilů a Johnsonova hierarchického klastrování s mírami adekvátnosti klastrů; zvolena řešení se čtyřmi pozicemi

Třídy (pozice) strukturně ekvivalentních uzlů v ukázkové síti

	CONCOR	Optimizace	Podobnosti profilů	Hierarchická CA
	8	4	8	8
	9	5	9	9
	7	6	2	15
	11	10	3	11
	10	12	1	10
	12	15	4	12
	13	14	7	13
	2	13	6	7
	3	16	5	6
	1	8	11	5
	4	2	10	3
	6	1	12	4
	5	3	13	2
	14	9	14	1
	15	7	15	14
klastry	16	11	16	16
Eta	0,623	0,35	0,481	0,35
Q	0,097	0,01	0,01	-0,055
E-I	-0,601	0,428	0,699	0,833
R2	0,363	0,759	N/A	N/A

stejně vazby o stejné intenzitě ke všem ostatním uzlům v síti. Uzly, o nichž je toto známo bez použití poziční analýzy pouze na základě měr centrality a koheze sítě, jsou v síti tři. Jsou jimi izolované uzly, které nemají žádné vazby neboli samé nuly v datové matici, díky čemuž mají perfektně korelované profily, a proto musí náležet do stejné ekvivalentní třídy. Jak je patrné z tabulky č. 1, tuto podmínku splňují všechna uvedená rozdělení vyjma hierarchické klastrové analýzy (complete link metoda), jejímž prostřednictvím byl uzel 15 zařazen do jiné pozice než 14 a 16. V řešení pomocí Tabu optimalizace jsou izoláty sice ve stejné pozici, avšak spolu s nimi je v ní ještě řada dalších uzlů. To by nemuselo nutně vadit, kdyby ovšem některé z těchto uzlů (např. 5 či 6) neměly výrazně vyšší centralitu než izoláty, což napovídá, že je z hlediska podobnosti vazeb tato pozice značně nehomogenní. V tomto řešení je také 4 zařazen do zvláštní pozice, aniž by se jeho profil jakkoliv odlišoval od některých jiných uzlů – naopak, velmi silně ($r > 0,95$) koreluje s profily uzlů 1, 2 a 3. Ve všech čtyřech navržených rozděleních se ve stejných pozicích objevují nejcentrálnější 8 a 9, dále 12 a 13 a společně také trojice uzlů 1, 2, 3. Z interpretačního

Table 1: Equivalent classes of nodes in the network according to the procedures of CONCOR, Tabu Search optimization, cluster analysis of profiles, and Johnson's Hierarchical Clustering, with cluster adequacy levels; four-class solutions have been selected

Classes (positions) of structurally equivalent nodes in the illustrative network

	CONCOR	Optimization	CA of profiles	Hierarchical CA
	8	4	8	8
	9	5	9	9
	7	6	2	15
	11	10	3	11
	10	12	1	10
	12	15	4	12
	13	14	7	13
	2	13	6	7
	3	16	5	6
	1	8	11	5
	4	2	10	3
	6	1	12	4
	5	3	13	2
	14	9	14	1
	15	7	15	14
clusters	16	11	16	16
Eta	0.623	0.35	0.481	0.35
Q	0.097	0.01	0.01	-0.055
E-I	-0.601	0.428	0.699	0.833
R2	0.363	0.759	N/A	N/A

hlediska se tedy jeví jako nejrozumnější řešení nabízená procedurou CONCOR a podobnostmi profilů (klastrovaných UCINETem přednastaveně metodou average link), protože v nich jsou navíc izolovaní aktéři sami v jedné pozici. Tento poněkud mikroskopický pohled je však dobré ještě doplnit mírami adekvátnosti klastrů a poté zvolit nejvhodnější reprezentaci struktury sítě.

Interpretaci blokmodelu lze v zásadě založit buď na atributech uzlů, popisu jednotlivých pozic, nebo pomocí image matice¹⁵ [Wasserman, Faust 1994: 408–417]. Vlastnosti uzlů, které lze využít k interpretaci blokmodelu, jsou k dispozici v podstatě dvě – míry centrality a pak také jejich atribuční data – tj. data užívaná v běžné statistice popisující vlastnosti jednotlivých uzlů, ať už se jedná o kardinální, nebo kategoriální proměnné. V tomto ohledu interpretaci také velice usnadňuje a zároveň i validizuje kvalitativní vzhled a znalost kontextu, který doprovází danou síť¹⁶. V první pozici jsou dva relativně okrajoví aktéři 12 a 13. Pro interpretaci by zde bylo vhodné nalézt atribut, který vhodně vysvětluje, proč spolu nespolupracují, ale s uzly v druhé pozici ano – může se např.

Tabulka 2: Do bloků uspořádaná matice s vyznačenými pozicemi podle CONCOR

Blokmodel procedury CONCOR se čtyřmi pozicemi

	12	13	9	11	8	7	10	4	5	2	1	3	6	14	15	16
12			1	1	2											
13			1	1	1											
9	1	1		3	10	3	1	3	1	5	3	5	2			
11	1	1	3		3	1				1	1	1				
8	2	1	10	3		3	1	3	2	5	3	4	2			
7			3	1	3			1	1	2	1	2	2			
10			1		1											
4			3		3	1			1	3	2	3	1			
5			1		2	1		1		1	1	1	1			
2			5	1	5	2		3	1		4	6	1			
1			3	1	3	1		2	1	4		4	1			
3			5	1	4	2		3	1	6	4		1			
6			2		2	2		1	1	1	1	1				
14																
15																
16																

Table 2: Blocked matrix with classes indicated according to CONCOR

CONCOR block model with four classes

	12	13	9	11	8	7	10	4	5	2	1	3	6	14	15	16
12			1	1	2											
13			1	1	1											
9	1	1		3	10	3	1	3	1	5	3	5	2			
11	1	1	3		3	1				1	1	1				
8	2	1	10	3		3	1	3	2	5	3	4	2			
7			3	1	3			1	1	2	1	2	2			
10			1		1											
4			3		3	1			1	3	2	3	1			
5			1		2	1		1		1	1	1	1			
2			5	1	5	2		3	1		4	6	1			
1			3	1	3	1		2	1	4		4	1			
3			5	1	4	2		3	1	6	4		1			
6			2		2	2		1	1	1	1	1				
14																
15																
16																

Tabulka 3: Do bloků uspořádaná matice s vyznačenými pozicemi podle klastrové analýzy podobností profilů

Blokmodel klastrové analýzy podobností profilů se čtyřmi pozicemi

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		4	4	2	1	1	1	3	3		1					
2	4		6	3	1	1	2	5	5		1					
3	4	6		3	1	1	2	4	5		1					
4	2	3	3		1	1	1	3	3							
5	1	1	1	1		1	1	2	1							
6	1	1	1	1	1		2	2	2							
7	1	2	2	1	1	2		3	3		1					
8	3	5	4	3	2	2	3		10	1	3	2	1			
9	3	5	5	3	1	2	3	10		1	3	1	1			
10								1	1							
11	1	1	1				1	3	3			1	1			
12								2	1		1					
13								1	1		1					
14																
15																
16																

Table 3: Blocked matrix with classes indicated according to cluster analysis of profiles

Cluster analysis of profiles block model with four classes

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		4	4	2	1	1	1	3	3		1					
2	4		6	3	1	1	2	5	5		1					
3	4	6		3	1	1	2	4	5		1					
4	2	3	3		1	1	1	3	3							
5	1	1	1	1		1	1	2	1							
6	1	1	1	1	1		2	2	2							
7	1	2	2	1	1	2		3	3		1					
8	3	5	4	3	2	2	3		10	1	3	2	1			
9	3	5	5	3	1	2	3	10		1	3	1	1			
10								1	1							
11	1	1	1				1	3	3			1	1			
12								2	1		1					
13								1	1		1					
14																
15																
16																

jednat o asistenty. V druhé pozici se nachází ústřední dvojice 8 a 9 spolu s 11, majícím poměrně vysokou mezilehlost, a dále 7 a 10. 10 se sice svými vazbami značně liší od ostatních členů této pozice, a proto by bylo nutné vysvětlit jeho zařazení třeba opět vhodným atributem, aby byl v tomto ohledu zařazen smysluplně. Až na něj se totiž jedná o aktéry centrální. Třetí pozice se skládá uzlů 1, 2, 3, 4, 5 a 6. Všichni jsou v síti poměrně významnými prvky, co se jejich centrality týče. Poslední pozici představují izolovaní aktéři, kteří v případě sítě spolupráce žádnou spolupráci nevykazují a pracují tak pouze o samotě.

Při pohledu na vazby jednotlivých pozic je vidět, že se v síti nacházejí dvě pozice (druhá a třetí), které mají husté vazby jak mezi sebou, tak i jedna k druhé. Pozice uzlů 1–6 nemá další vazby, zatímco pozice druhá má navíc vazbu na pozici první, které samy mezi sebou vazbu nemají, ale vážou se k pozici aktérů centrálních. Takto by se dalo říci, že by možná uzlu 10 více slušelo zařazení do první pozice, neboť na její aktéry nemá žádné vazby a sám se váže jen k 8 a 9, tj. aktérům pozice druhé. Celkově se tedy síť skládá z centrální třídy hustě a často spolupracujících uzlů (7, 8, 9, 10, 11 – podotýkám, že zařazení 10 do této pozice je problematické), kteří čteně interagují jak mezi sebou, tak i s dalšími pozicemi, zejména s pozicemi uzlů 1–6. Uzly 1–6 spolupracují pouze s ústředními nebo jen v rámci své pozice a vyznačují se středními až vyššími hodnotami měř centrality. Další pozici tvoří pouze dva aktéři – asistenti. Poslední pozici tvoří izolovaní aktéři. Tento obecný model sítě ve tvaru (1) ústřední a dobře propojení, (2) méně centrální a dobře propojení, (3) asistující a (4) izolovaní s mírnými odlišnostmi ukazuje i blokmodel podle klastrové analýzy podobností profilů (tabulka č. 3). Tabulka č. 4 představuje základ pro vytvoření image matice a následně i image grafu (tabulka č. 5 a graf č. 3), v nichž byly 1- a 0-bloky zvoleny na základě kritéria maximální hodnoty, kde je touto hodnotou alespoň nějaká vazba, tedy 1.

Závěr

V předloženém textu byl představen koncept ekvivalence a s ním související pojetí pojmů role a pozice v rámci analýzy sítí. Pro dva nejvíce užívané typy ekvivalence, strukturální a regulární, byly také představeny různé způsoby, jak v sítích identifikovat uskupení vzájemně si ekvivalentních uzlů, a to se zvláštním důrazem na blokové modelování. Díky konceptu ekvivalence lze síť zredukovat na uskupení uzlů (pozice) a vazby mezi těmito pozicemi (role), čímž vzniká zjednodušená reprezentace sítě, která přitom ale zachovává její podstatné strukturální vlastnosti. To umožňuje síť mezi sebou porovnávat, teoretizovat o struktuře dané sítě nebo prostě zjednodušit velkou a komplikovanou síť. Je však potřeba mít vždy na paměti nedostatky spojené s explorativními metodami a opírat blokované modelování a jeho alternativy o znalost teorie. Dobrou praxí je také zkoušet a vzájemně porovnávat vícero metod.

Přestože blokové modelování a obecně analýza pozic momentálně na špičce výzkumu v oblasti sítí (ať už sociálních, nebo v novém oboru síťové vědy) poněkud ustoupily statistickému modelování konfigurací v sítích (např. dyády či různé druhy triád), jedná se stále o oblast v SNA, která nejenže má své místo, ale i potenciál k dalšímu rozvoji [Robins 2015: 189]. V předložené stati jsem se pokusil nastínit možnosti, ale i omezení metod identifikace ekvivalentních tříd a demonstrovat základy jejich aplikace. Z dalších empiricky zdařilých aplikací, které mohou sloužit jako exemplární ukázky poziční analýzy, lze uvést historicky laděné studie Padgetta a Ansell [1993] o florentském rodu Medicejských, biomedicínský výzkum Breigera [1976], zkoumání soudobého světosystému [Kick et al. 2011], protestních hnutí od Bearmana a Everetta [1993] nebo instruktivní článek o hudebních festivalech ve Velké Británii Crossleyho a Emmsově [2016]. Tyto a mnohé další práce dokládají, že blokové modelování a potažmo SNA jako celek mají v sociálních vědách své místo a účel.

Tabulka 4: Matice hustoty bloků se čtyřmi pozicemi podle CONCOR

Matice hustoty bloků podle CONCOR

	1	2	3	4
1	0	0,7	0	0
2	0,7	2,5	1,7	0
3	0	1,7	2	0
4	0	0	0	0

Table 4: Block density matrix with four classes, CONCOR

Block density matrix, CONCOR

	1	2	3	4
1	0	0,7	0	0
2	0,7	2,5	1,7	0
3	0	1,7	2	0
4	0	0	0	0

Tabulka 5: Image matice podle CONCOR s kritériem maximální hodnoty (= 1)

Image matice podle CONCOR

	1	2	3	4
1	0	1	0	0
2	1	1	1	0
3	0	1	1	0
4	0	0	0	0

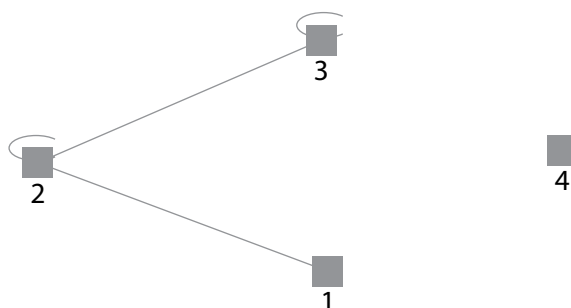
Table 5: Image of CONCOR matrix with maximum value criterion (= 1)

Image of CONCOR matrix

	1	2	3	4
1	0	1	0	0
2	1	1	1	0
3	0	1	1	0
4	0	0	0	0

Graf 3: Image graf čtyř pozic podle CONCOR

Figure 3: Image graph of four classes, CONCOR



literatura

- Batagelj, Vladimir, Patrick Doreian, Anuska Ferligoj. 2011. Positions and Roles. Pp 434–447 in Scott, John, Peter J. Carrington (eds.). *Handbook of Social Network Analysis*. Thousand Oaks: SAGE.
- Bearman, Peter S., Kevin D. Everett. 1993. „The Structure of Social Protest, 1961–1983.“ *Social Networks* 15 (2): 171–200, [http://dx.doi.org/10.1016/0378-8733\(93\)90004-5](http://dx.doi.org/10.1016/0378-8733(93)90004-5).
- Bellotti, Elisa. 2014. *Qualitative Networks*. London: Routledge.
- Borgatti, Steve P., Martin G. Everett, Linton C. Freeman. 2002. *UCINET 6 for Windows: Software for Social Network Analysis*. Harvard: Analytic Technologies.
- Borgatti Steve P., Martin G. Everett, Jeffrey C. Johnson. 2013. *Analyzing Social Networks*. London: SAGE.
- Breiger, Ronald L. 1976. „Career Attributes and Network Structure: A Blockmodel Study of Biomedical Research Specialty.“ *American Sociological Review* 41: 117–135.
- Scott, John, Peter Carrington (eds.). 2011. *Handbook of Social Network Analysis*. Thousand Oaks: SAGE.
- Císař, Ondřej, Jiří Navrátil. 2014. „Promoting Competition or Co-operation? The Impact of EU Funding on Czech Advocacy Organizations.“ *Democratization* 22 (3): 1–24, <http://dx.doi.org/10.1080/13510347.2013.869742>.
- Crossley, Nick., Rachel. Emms. 2016. „Mapping the Musical Universe: A Blockmodel of UK Music Festivals, 2011–2013.“ *Methodological Innovations* 9 (1): 1–14, <http://dx.doi.org/10.1177/2059799116630663>.
- Diviák, Tomáš 2015. *Pavučiny zločinu: Korupce v perspektivě analýzy sociálních sítí*. Univerzita Karlova v Praze: Nepublikovaná diplomová práce.
- Domínguez, Silvia, Betina Hollstein. 2014. *Mixed Methods Social Networks Research*. Cambridge: Cambridge University Press.
- Doreian, Patrick, Vladimir Batagelj, Anuska Ferligoj. 2005. *Generalized Blockmodeling*. Cambridge: Cambridge University Press.
- Everton, Sean F. 2012. *Disrupting Dark Networks*. Cambridge: Cambridge University Press.
- Granovetter, Mark. 1973. The Strength of Weak Ties. *American Journal of Sociology* 78 (6): 1360–1380.
- Hanneman, Robert A., Mark Riddle. 2005. *Introduction to Social Network Methods* [online]. Riverside: University of California [cit. 6. 3. 2016]. Dostupné z: <http://faculty.ucr.edu/~hanneman/nettext/>.
- Hendl, Jan. 2009. *Přehled statistických metod*. Praha: Portál.

- Kick, Edward L., Laura A. McKinney, Steve McDonald, Andrew Jorgenson. 2011. „A Multiple-Network Analysis of the World System, 1995–1999.“ Pp. 311–329 in Peter J. Carrington, John Scott (eds.). *Handbook of Social Network Analysis*. Thousand Oaks: SAGE.
- Kolaczyk, Eric D. 2009. *Statistical Analysis of Network Data*. New York: Springer.
- Lorrain, François, Harrison C. White. 1971. „Structural Equivalence of Individuals in Social Networks.“ *Journal of Mathematical Sociology* 1 (1): 49–80, <http://dx.doi.org/10.1080/0022250X.1971.9989788>.
- Mazák, Jaromír, Pavla Homolová, Tomáš Diviák, Romana Imříšková, Kristýna Chábová, Martina Krynská, Viktorie Kolínská. 2015. *Využití analýzy sociálních sítí ve vyšetřování*. Kompetenční centrum IBM při katedře sociologie Filozofické fakulty Univerzity Karlovy v Praze.
- Morselli, Carlo. 2009. *Inside Criminal Networks*. New York: Springer.
- Mrvar, Andrej, Wouter de Nooy, Vladimir Batagelj. 2005. *Exploratory Social Network Analysis with Pajek*. Cambridge: Cambridge University Press.
- Padgett, John. F., Christopher Ansell. 1993. „Robust Action and the Rise of the Medici, 1400–1434.“ *American Journal of Sociology* 93 (6): 1259–1319, <http://dx.doi.org/10.1086/230190>.
- Petrusek, Miloslav. 1969. *Sociometrie*. Praha: Svoboda.
- Robins, Garry. 2015. *Doing Social Network Research*. London: SAGE.
- Schmidt, Jan, Jiří Šubrt. 2010. „Analýza sociálních sítí.“ Pp. 332–364 in Šubrt, Jiří. et al. *Soudobá sociologie IV (Aktuální a každodenní)*. Praha: Karolinum.
- Vašát, Petr, Josef Bernard. 2015. „Formování komunit nebo sociální integrace? Analýza personálních sítí ukrajinských imigrantů v Plzni.“ *Sociologický časopis* 51 (2): 199–225, <http://dx.doi.org/10.13060/00380288.2015.51.2.159>.
- Wassermann, Stanley, Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.
- White, Douglas R., Karl P. Reitz. 1983. „Graph and Semigroup Homomorphisms on Networks and Relations.“ *Social Networks* 5: 193–234., [http://dx.doi.org/10.1016/0378-8733\(83\)90025-4](http://dx.doi.org/10.1016/0378-8733(83)90025-4).
- Žiberna, Aleš. 2007. „Generalized Blockmodeling of Valued Networks.“ *Social Networks* 29: 105–126, <http://dx.doi.org/10.1016/j.socnet.2006.04.002>.
- Žiberna, Aleš. 2008. „Direct and Indirect Approaches to Blockmodeling of Valued Networks in Terms of Regular Equivalence.“ *The Journal of Mathematical Sociology* 32: 57–84, <http://dx.doi.org/10.1080/00222500701790207>.
- Žiberna, Aleš. 2009. „Evaluation of Direct and Indirect Blockmodeling of Regular Equivalence in Valued Networks by Simulations.“ *Metodološki Svezki* 6: 99–134.

Mgr. Tomáš Diviák je výzkumným pracovníkem katedry sociologie Filozofické fakulty Univerzity Hradec Králové. Je také doktorandem katedry sociologie Filozofické fakulty Univerzity Karlovy a Inter-university Center for Social Science Theory and Methodology, University of Groningen, Nizozemí. Ve svém výzkumu i výuce se zabývá analytickou sociologií a metodologií analýzy sociálních sítí a jejími aplikacemi v oblastech studií organizovaného zločinu, veřejného mínění nebo politiky.

Lze jej kontaktovat na adrese: tomas.diviak@seznam.cz

poznámky

¹ Základ této stati vychází z metodologické části autorovy diplomové práce [Diviák 2015].

² Ke vzniku tohoto článku svými radami a připomínkami přispěl Tom A. B. Snijders, kterému tímto děkuji. Díky patří také oběma anonymním recenzentům, jejichž postřehy a připomínky napomohly k vylepšení tohoto textu.

³ Z anglického social network analysis.

⁴ V případě dvou odlišných druhů uzlů v jedné síti, které mohou navazovat vazby jen napříč druhy (např. jedinci a organizace, kde je vazba „být členem“, tj. jedinec může mít vazbu k organizaci a ne k jinému jedinci), hovoříme o specifickém případě tzv. bimodálních (two-mode) sítí.

⁵ Česká terminologie pro SNA ještě zdaleka není ustálená, proto zde uvádím i anglické předlohy daných pojmů.

⁶ Existuje řada přibližně synonymně užívaných pojmů jako např. community či shluky, ale vzhledem k tomu, že tyto termíny mají jiný význam mimo SNA, uchyluji se zde k užívání termínu podskupina.

⁷ Úvod do SNA prostřednictvím Pajeku pak nabízí text Mrvara, de Nooye a Batagelje [2005].

⁸ Uskupení ekvivalentních uzlů bývají také označovány jako ekvivalentní třídy nebo pozice [Wasserman, Faust 1994: 357].

⁹ Nicméně v případě, že by se mezi sebou učitelé vzájemně učili (tj. měli mezi sebou vazby „učit někoho“), pak by mohli být identifikováni nejen jako pozice, ale i jako podskupina.

¹⁰ Další někdy uváděné typy ekvivalence jsou např. automorfni (založená na matematické vlastnosti grafů zvané isomorfismus) a stochastická (založená na pravděpodobnosti výskytu vazeb), nicméně třeba Borgatti, Everett a Johnson [2013] je ani nezmiňují ve své knize. Jediná mně známá práce dostatečně detailně popisující oba tyto typy je starší úvod od Wassermana a Faustové [1994]. Reálně jsou ve výzkumu v porovnání s regulární a strukturální ekvivalencí její jiné definice marginální [Everett, osobní komunikace].

¹¹ Jedná se o síť s jednou množinou uzlů a větším počtem relací než jedním. Příkladem jsou třeba síť přátelství, nepřátelství nebo spolupráce ve školních třídě nebo vztahy obchodní, diplomatické a vojenské mezi státy.

¹² Míry podobnosti, např. míru strukturální ekvivalence dvou uzlů, lze mimo užití popsané v tomto článku využít i jako dyadické nezávisle i závisle proměnné v QAP regresních modelech, které umějí zahrnout vzájemnou interdependenci objektů a relační data [Borgatti et al. 2013: 220]. Při jakékoliv statistické analýze sítí je však nutné mít v patrnosti to, že nelze využít klasickou statistickou inferenci, neboť relační data inherentně narušují předpoklad nezávislosti pozorování. Alternativ je celá řada, blíže se o nich lze dočíst ve výše zmíněných publikacích či v Kolarczyk [2009] či Scott, Carrington [2011].

¹³ Z anglického „CONvergent CORrelations“.

¹⁴ V angličtině se používají poněkud elegantnější termíny sink (přijímač) a source (vysílač).

¹⁵ Ačkoliv jsou image matice standardně užívány ke zjednodušenému zachycení blokmodelů, zde uvádím celé do bloků permutované matice, abych se mohl odkazovat při interpretaci na konkrétní uzly či vazby a poukázat tak na chyby a další aspekty jednotlivých řešení. To by image matice neumožnily. Věřím, že u takto malých sítí jsou celé matice stále přehledné.

¹⁶ Komplementarita SNA a kvalitativních metod je mimochodem v současnosti hojně diskutovaným tématem [viz např. Bellotti 2014; Domínguez, Hollstein 2014].