Sveučilište u Zagrebu
Prirodoslovno–matematički fakultet

Hermina Petric Maretić

Dinamička analiza mreža – evolucija istraživačkog područja temeljem radova objavljenih u sklopu DESIGN konferencije 2002–2014

Zagreb, 2014.

Ovaj rad izrađen je na Matematičkom odsjeku Prirodoslovno–matematičkog fakulteta u Zagrebu, pod mentorstvom prof. dr. sc. Maria Štorge s Fakulteta strojarstva i brodogradnje i komentorstvom prof. dr. sc. Saše Singera s Matematičkog odsjeka Prirodoslovno–matematičkog fakulteta, te je predan na natječaj za dodjelu Rektorove nagrade u akademskoj godini 2013./2014.

Sadržaj

1 Uvod

Napretkom društva i pojavom velikih količina informacija prirodno se pojavila potreba za povezivanjem podataka i analizom informacija. Osnovni koncept, ali i intuitivno prva asocijacija za povezivanje su mreže. Pojavom mreža složenih struktura uzrokovanih kompleksnošću podataka koje je potrebno povezati i vezama čija protočnost mora podržavati širok aspekt struktura komunikacije, pojavila se i potreba za strukturiranjem i analizom mreža.

Što je zapravo kompleksna mreža i kako je prepoznati? Bitno je uočiti razliku između kompleksnog i kompliciranog sustava − avion je vrlo kompliciran, ali nije kompleksan. Iako ne postoji prihvaćena definicija kompleksnog sustava, možemo ih prepoznati po nekim obilježjima. Pogledajmo primjer Interneta. Za njega ne postoji globalni plan za strukturu mreže, već se ona razvija samostalno. S druge strane, avion ima predefiniran plan gradnje. Internet se mijenja "iznutra", odnosno svatko može dodati čvor i tako promijeniti mrežu koja s vremenom evoluira. Usto, kompleksne mreže često imaju sposobnost prilagođavanja, pa mogu i odolijevati raznim napadima. Jasno je da za avion s nekoliko kvarova ne možemo reći da funkcionira na zadovoljavajućoj razini, dok u slučaju kvara nekoliko rutera, Internet i dalje dobro funkcionira na globalnoj razini.

1.1 Važnost mreža

Danas gotovo ne postoji područje znanosti koje tokom analize ili konstrukcije određenih koncepata ne koristi neke oblike mreže. Primjer jedne izuzetno složene mreže možemo dobiti proučavanjem strukture i funkcioniranja mozga. Takve mreže zovu se strukturalno funkcionalne mreže [1]. Pojam mrežne strukture prirodno se pojavljuje u proučavanju socijalnih odnosa i povezanosti. Ovakve vrste mreža vrlo su bitne za razvoj socijalnih odnosa, politike, ekonomije i bankarstva.

Primjene te vrste mreža možemo pronaći čak i u modeliranju kretanja dionica [2]. Osim u znanstvene svrhe, mreže se pojavljuju i kao strukture u arhitekturi poput telekomunikacijskih mreža [4], mreža prometnica [5], interneta, urbanih mreža i mnogih drugih.

1.2 DESIGN konferencije

Konferencije iz serije DESIGN održavaju se s prekidima preko 30 godina prateći razvoj područja koje se naziva znanost o konstruiranju (eng. Design Science). U početnim godinama konferencije iz serije DESIGN uglavnom su bile događaji s domaćim predznakom, uz pokojeg sudionika iz inozemstva.

Od 2002., uz podršku novo-osnovanog međunarodnog udruženja The Design Society (`http://www.designsociety.org`), DESIGN postaje kvalitativno i kvantitativno najveća europska konferencija u području koja redovno okuplja oko 300 sudionika iz 30-tak zemalja, koji prosječno tijekom konferencije prezentiraju 200 znanstvenih radova. Kvaliteta konferencije je vidljiva iz sastava programskog i znanstvenog odbora u kojem sudjeluje preko 200 vrhunskih stručnjaka u području. DESIGN danas predstavlja ključni događaj u zajednici, na kojem se okupljaju sudionici iz industrije i akademije kako bi predstavili i razmijenili iskustva vezana uz suvremeni razvoj tehničkih sustava i procesa, pokušavajući dati odgovore na suvremene zahtjeve društva vezane uz održivost, inovativnost, globalizaciju, te socio-tehničke aspekte razvoja.

Zbog bogate povijesti i tradicije, te međunarodne prepoznatljivosti, konferencije iz serije DESIGN reflektiraju razvoj ovog relativno mladog znanstvenog područja koje se počelo intenzivnije razvijati na europskim sveučilištima u 90-tim godinama 20. stoljeća. Promjena tematika i fokusa radova objavljenih na DESIGN konferencijama reflektira promjene koje su se događale u istraživačkim grupama u području. Analiza te promjene objašnjava kako se znanost o konstruiranju, prije fokusirana na oblikovanje komponenti proizvoda i metodički proces konstruiranja, transformirala u inter- i trans-disciplinarno područje istraživanja razvojnih procesa potrebnih za stvaranje kompleksnih tehničkih sustava u globalnom okruženju, upravljanja kompleksnim sustavima i procesima, interakcije tehničkih sustava s okolinom, te ponašanja ljudi u tom razvoju s fokusom na kreativne i inovativne procese.

2 Ciljevi rada

Osnovni ciljevi ovog rada su analiza kompleksne mreže na statičkoj i dinamičkoj razini te uspoređivanje mreže s različitim teorijskim modelima mreža. Konkretno, proučavat ćemo skup članaka na konferencijama, te ključne riječi koje se pojavljuju u njima. Na temelju tih podataka, želimo za svaku konferenciju modelirati statičku mrežu koja će je opisivati. Te mreže ćemo promatrati kroz godine i tako dobiti dinamičku mrežu na kojoj želimo ispitati različita svojstva u nadi da ćemo dobiti bolji uvid u strukturu konferencija. Cilj je dinamičkom analizom utvrditi razvoj znanosti kroz godine. Pokušat ćemo izdvojiti riječi koje su u pojedinim godinama bile od iznimnog značenja te promatrati njihovo ponašanje u ostalim godinama.

Osim toga, želimo upoznati osnovne modele mreža i s njima usporediti promatranu mrežu. Cilj je uočiti sličnosti i razlike s postojećim modelima te prepoznati bitna svojstva dane mreže kako bismo dobili bolji uvid u strukturu mreže. Naime, poznavanje strukture i svojstava mreže koje stječemo navedenim usporedbama moglo bi nam pomoći da bolje razumijemo ponašanje mreže i promjene u njoj, a time nam dati i novi, jasniji uvid u povezanost tematika u znanosti o konstruiranju te promjene koje mogu opisati njezin razvoj.

3 Reprezentacija i analiza mreža

Teorija koja opisuje i analizira kompleksne mreže svoje temelje ima u teoriji grafova i statističkoj analizi. Kompleksnu mrežu opisujemo pomoću grafa G, tako da podatke reprezentiramo čvorovima, a veze između njih bridovima.

Ovisno o svojstvima mreže, taj graf može biti usmjeren ili neusmjeren, težinski ili bez težina. Stoga u ovom poglavlju iznosimo osnovne pojmove teorije grafova, kao i neke specijalne veličine bitne za analizu mreža.

3.1 Osnovni pojmovi teorije grafova[12]

Definicija 3.1 ***(Neusmjereni) Graf*** *je uređeni par* $G = (V, E)$*, gdje je* $V = V(G)$ *neprazan skup čije elemente nazivamo vrhovima, a* $E = E(G)$ *je familija dvočlanih podskupova od* V *koje nazivamo bridovima.*

Pritom vrhove u *i* v *koji su pridruženi bridu* e *nazivamo krajevima brida* e*, a brid čiji se krajevi podudaraju nazivamo* ***petljom****.*

Definicija 3.2 ***Usmjereni graf*** *je uređeni par* $G = (V, E)$*, gdje je* $V = V(G)$ *neprazan skup čije elemente nazivamo vrhovima, a E=E(G) je skup uređenih parova elemenata iz* V*, čije elemente nazivamo usmjerenim bridovima.*

Definicija 3.3 ***Multiskup*** M *na skupu* S *je uređeni par* (S, m)*, gdje je* m *funkcija,* $m : S \to \mathbb{N}_0$*.*

Definicija 3.4 ***Neusmjereni (Usmjereni) multigraf*** *je neusmjereni (usmjereni) graf čiji bridovi čine multiskup.*

Napomena 3.5 *Ukoliko želimo naglasiti da se ne radi o multigrafu, graf nazivamo* ***jednostavnim grafom****.*

Definicija 3.6 ***Težinski graf*** *G je graf kojem je pridružena funkcija $f : E \to \mathbb{R}$, koja svakom bridu iz grafa pridružuje njegovu težinu. Tu funkciju nazivamo* ***težinskom funkcijom****.*

Napomena 3.7 *Graf koji nije težinski možemo nazivati još i bestežinskim grafom.*

Definicija 3.8 *Za par vrhova u i v kažemo da su* ***susjedni*** *ako postoji brid e kojemu su oni krajevi. Pritom kažemo da je brid e* ***incidentan*** *s vrhovima u i v te u slučaju usmjerenog grafa pišemo $e = (u, v)$, a u slučaju neusmjerenog $e = \{u, v\}$. U općenitom slučaju koristi se $e = uv$.*

Definicija 3.9 ***Stupanj vrha*** *v, u oznaci $\deg(v)$, u neusmjerenom (multi)grafu G je broj bridova grafa G incidentnih s v, pri čemu se svaka petlja računa kao dva brida.*

Izolirani vrh *je vrh stupnja 0.* ***Ulazni stupanj vrha*** *v u usmjerenom (multi)grafu G je broj bridova grafa G koji završavaju u vrhu v, odnosno broj bridova oblika $e = (x, v)$.*

Analogno se definira ***izlazni stupanj vrha****.*

Definicija 3.10 ***Šetnja*** *u (usmjerenom) (multi)grafu je niz $(v_0, e_1, v_1, e_2, v_2, \dots, e_k, v_k)$, gdje je e_i brid $v_{i-1}v_i$, za $i = 1, \dots, k$.* ***Put*** *je šetnja u kojoj su svi vrhovi različiti (osim eventualno početnog i završnog).* ***Duljina puta*** *je broj bridova u nizu za bestežinske (multi)grafove, odnosno suma vrijednosti težinske funkcije svih bridove u nizu za težinske (multi)grafove.*

Kažemo da su dva vrha ***povezana*** *ako postoji put između njih.*

Kažemo da je graf ***povezan*** *ako su svi njegovi vrhovi povezani.*

Definicija 3.11 ***Komponenta povezanosti*** *(usmjerenog) (multi)grafa je maksimalni povezani podgraf grafa.*

Definicija 3.12 *Neka je $G = (V, E)$ (usmjereni) (multi)graf i $n = |V|$ broj vrhova. Neka je $V = \{v_1, v_2, \ldots, v_n\}$.* ***Matrica susjedstva*** *grafa G je matrica $A \in M^{n \times n}$ koja na mjestu (i, j) ima vrijednost jednaku broju bridova incidentnih vrhovima v_i i v_j.*

Napomena 3.13 *Primijetimo da u slučaju jednostavnog grafa matrica A poprima vrijednosti u skupu $\{0, 1\}$.*

3.2 Neke veličine bitne za analizu mreža

U analizi mreže vrlo je bitno moći odrediti važnost određenog vrha te istaknuti one koji su najutjecajniji. U tu svrhu definiramo nekoliko veličina, tzv. centralnosti, među kojima su najpopularnije centralnost stupnja (eng. degree centrality), centralnost blizine (eng. closeness centrality) i centralnost između vrhova (eng. betweenness centrality).

Napomena 3.14 *U daljnjem tekstu sve su definicije izražene za neusmjerene grafove bez težina, ukoliko nije drugačije naglašeno. Definicije se proširuju na težinske i/ili usmjerene grafove na način analogan gornjim primjerima.*

U neusmjerenom grafu **centralnost stupnja** vrha i jednaka je stupnju tog vrha, $k_i = \deg(v)$.

U usmjerenom grafu definiramo **ulaznu centralnost stupnja** $k_{in,i}$, **izlaznu centralnost stupnja** $k_{out,i}$ i **ukupnu centralnost stupnja** $k_i = k_{in,i} + k_{out,i}$.

Napomena 3.15 *Često se kao centralnost stupnja koristi i odnos stupnja vrha s najvećim mogućim stupnjem vrha,*

$$k_i = \frac{\deg(v)}{n-1}.$$

Ova veličina na najjednostavniji način opisuje koliko je dobro vrh i povezan s ostalim vrhovima u grafu.

Definicija 3.16 *Neka je $G = (V, E)$ graf i označimo s $l_{i,j}$ duljinu najkraćeg puta od vrha i do vrha j. Tada*

$$d_G = \max_{i,j} l_{i,j}$$

*zovemo **promjer grafa** G.*

Ako je još $n = |V|$ broj vrhova grafa G, tada

$$\langle l \rangle = \frac{1}{n(n-1)} \sum_{i,j} l_{i,j}$$

*označava **prosječnu duljinu najkraćeg puta**.*

Napomena 3.17 *Ukoliko ne postoji put između vrha i i vrha j, formalno definiramo duljinu najkraćeg puta kao $l_{i,j} = \infty$. Ipak, često želimo izračunati neka svojstva i u grafovima koji nisu povezani, pa za tu udaljenost koristimo neku veliku konstantu ili broj vrhova u grafu.*

Definicija 3.18 *Neka je $G = (V, E)$ graf te $i \in V$ neki njegov vrh. Za vrh $j \in V$ označimo s $l_{i,j}$ duljinu najkraćeg puta od vrha i do vrha j. Tada sa*

$$g_i = \frac{1}{\sum_{j \neq i} l_{i,j}}$$

*definiramo **centralnost blizine**.*

Velika centralnost blizine upućuje na vrh čiji je prosječni najkraći put do drugih čvorova malen, odnosno baš na vrh koji je blizu ostalim čvorovima.

Dosad nabrojene veličine dobro opisuju povezanost vrha s ostatkom grafa, ali mogu zanemariti vrhove koji nemaju veliki stupanj, a povezuju razna, inače slabo povezana područja. Centralnost između vrhova opisuje važnost vrha u ulozi mosta između različitih dijelova mreže.

Definicija 3.19 *Neka je σ_{hj} broj najkraćih puteva od h do j, a $\sigma_{hj}(i)$ broj najkraćih puteva od h do j koji prolaze kroz i. Tada definiramo*

$$L_i = \sum_{h \neq j \neq i} \sigma_{hj}(i)$$

naglašenu centralnost *(eng. stress centrality) i*

$$b_i = \sum_{h \neq j \neq i} \frac{\sigma_{hj}(i)}{\sigma_{hj}}$$

centralnost između vrhova*.*

Kad bismo promatrali utjecajnost nekog elementa u mreži, vrlo važan faktor bio bi povezanost s drugim utjecajnim elementima mreže. Matematički to možemo zapisati pomoću sustava jednadžbi

$$x_v = \frac{1}{\lambda} \sum_{t \in G} a_{v,t} x_t,$$

gdje je $A = (a_{v,t})$ matrica susjedstva, t i v vrhovi grafa $G = (V, E)$, a λ neka konstanta. Zapišemo li taj sustav matrično, vidimo da se zapravo radi o svojstvenom problemu:

$$Ax = \lambda x.$$

Uz dodatni uvjet da sve vrijednosti u x budu pozitivne, može se pokazati da traženim svojstvima rezultira samo svojstveni vektor uz najveću svojstvenu vrijednost [14].

Definicija 3.20 *Neka je $G = (V, E)$ graf i $A = (a_{v,t})$ matrica susjedstva. Tada je* ***centralnost svojstvenog vektora*** *jednaka svojstvenom vektoru pridruženom najvećoj svojstvenoj vrijednosti λ matrice A.*

U kompleksnim mrežama jedna od vrlo bitnih stvari je sklonost formiranju jako povezanih grupa, tzv. klastera.

Definicija 3.21 *Neka je t_i broj veza među susjedima vrha i te neka je k_i stupanj vrha i. Tada* ***koeficijent klasteriranja*** *(eng. clustering cofficient) definiramo kao*

$$C(i) = \frac{t_i}{k_i(k_i - 1)/2}$$

za $k_i > 1$. Za $k_i \leq 1$ definiramo $C(i) \equiv 0$. Na razini cijele mreže definiramo ***prosječni koeficijent klasteriranja*** *(eng. average clustering coefficient) kao*

$$\langle C \rangle = \frac{1}{N} \sum_i C(i).$$

Definicija 3.22 *Definiramo funkciju distribucije $\mathrm{P}(k)$ kao vjerojatnost da je nasumično odabran vrh stupnja k. Ta se distribucija zove* ***distribucija stupnja*** *(eng. Degree distribution). U slučaju usmjerenog grafa, definiramo dvije distribucije $\mathrm{P}(k_{in})$ i $\mathrm{P}(k_{out})$.*

Srednji stupanj *mreže je*

$$\langle k \rangle = \frac{1}{N} \sum_i k_i = \sum_k k\,\mathrm{P}(k) = \frac{2E}{N}.$$

Analogno definiramo srednji ulazni i izlazni stupanj (primijetimo da moraju biti jednaki) za usmjerene grafove.

*Tada je n-**ti moment distribucije***

$$\langle k^n \rangle = \sum_k k^n \, \mathrm{P}(k).$$

Na analogan način definiramo i **distribuciju centralnosti između vrhova**, te **prosječnu centralnost između vrhova**.

3.3 Općenita svojstva mreža

Sociološki pokus iz 1967. inspirirao je znanstvenike na promatranje još jednog svojstva u kompleksnim mrežama. Nasumično odabranim ljudima u SAD-u podijeljena su pisma adresirana na neke druge, također nasumično odabrane, ljude u SAD-u. Cilj je bio dostaviti pismo do primatelja, ali uz pravilo da se pismo smije prenositi samo preko poznanika. Eksperiment je pokazao da je potreban vrlo malen broj poznanika da bi pismo dostiglo cilj, u prosjeku njih 6. Ipak, treba napomenuti da je eksperiment uspio samo u $\approx 20\%$ slučajeva, u $\approx 80\%$ slučajeva pismo nikad nije stigno na destinaciju [6]. Ovo svojstvo u kompleksnim mrežama nazvat ćemo svojstvo malog svijeta (eng. small-world property). Mreža koja ima ovo svojstvo mora za prosječnu duljinu najkraćeg puta imati neki mali broj. Ipak, logično je da on smije rasti s veličinom mreže.

Definicija 3.23 *Neka je $\langle l \rangle$ prosječna duljina najkraćeg puta. Za kompleksnu mrežu kažemo da ima **svojstvo malog svijeta** ako*

$$\langle l \rangle \backsim O(\log(N)).[11]$$

Oblik funkcija distribucija koje opisuju svojstva mreže dijeli mreže na dvije velike klase. U klasu **statistički homogenih mreža** spadaju sve mreže čije distribucije stupnja i distribucije centralnosti između vrhova imaju "lagane repove" poput normalne ili Poissonove distribucije. **Statistički heterogene mreže** karakterizirane su iskrivljenim (eng. skewed) distribucijama s "teškim repovima", najčešće distribucijama zakona potencija (eng. power law distribution) [8].

U homogenim mrežama sva se svojstva ponašaju onako kako se na prvi pogled i očekuje od mreže. Tako se većina statističkih svojstava promatranih na razini vrhova grupira oko očekivanih vrijednosti tih svojstava. Funkcija distribucije stupnja postiže maksimum u očekivanju i "brzo pada", što nam govori da će najveći broj vrhova imati stupanj jednak očekivanom stupnju, odnosno mod će biti vrlo blizak očekivanju. S druge strane, u heterogenim mrežama točka očekivanja nije istaknuta točka u funkciji distribucije.

Također, pri otkrivanju homogenih mreža koristi se činjenica da u njima varijanca nije velika.

Definicija 3.24 ***Koeficijent heterogenosti*** *definira se kao*

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle},$$

gdje $\langle k \rangle$ *predstavlja prvi moment (očekivanje) distribucije stupnja, a* $\langle k^2 \rangle$ *drugi moment distribucije stupnja.*

Mreža se smatra homogenom ako vrijedi $\kappa \backsim \langle k \rangle$, a heterogenom u slučaju $\kappa \to \infty$. U praksi ćemo mrežu smatrati heterogenom ako vrijedi $\kappa \gg \langle k \rangle$.

4 Modeliranje mreža

Prvi korak u boljem razumijevanju strukture mreže i predviđanju ponašanja mreže je aproksimiranje mreže nekim teoretskim modelom. U nastavku navodimo nekoliko najpoznatijih i najkorištenijih modela mreže.

Napomena 4.1 *U ovom poglavlju radi jednostavnosti promatrat ćemo neusmjerene grafove bez težina.*

4.1 Modeli mreža

Među najjednostavnijim modelima mreža koji uključuju stohastičnost definitivno je **Erdős–Renyi** model. Model predlaže dvije vrlo slične metode stvaranja mreže:

1. Neka je n broj vrhova i $|E|$ broj bridova grafa kojeg želimo modelirati. Graf $G_{n,|E|}$ konstruiramo tako da za svaki od $|E|$ bridova na nasumičan način odaberemo krajeve iz skupa od n vrhova.

2. Neka je n broj vrhova i $p \in [0, 1]$ vjerojatnost povezivanja dvaju vrhova. Tada graf $G_{n,p}$ konstruiramo tako da svaki od $\frac{n(n-1)}{2}$ bridova generiramo s vjerojatnošću p, odnosno ne generiramo s vjerojatnošću $1 - p$.

Prva metoda opisuje izvornu formulaciju Erdős–Renyi grafa iz 1959. godine, a druga jednostavnu preinaku modela koju je iste godine predložio Gilbert. Zbog jednostavnosti, u ovom radu koristit ćemo Gilbertovu metodu [13, 16].

Broj bridova u grafu generiranom na ovaj način u prosjeku iznosi $\langle |E| \rangle = \frac{n(n-1)}{2}p$. S obzirom na to da svaki brid pridonosi stupnju dvaju vrhova, tada prosječni stupanj brida u grafu iznosi

$$\langle k \rangle = \frac{2 \langle |E| \rangle}{n} = (n-1)p \simeq np.$$

Dok mreže u stvarnom životu, unatoč rastu, najčešće imaju konačan i ograničen srednji stupanj, vidi se da taj stupanj u ovom modelu divergira kako n raste. Iz tog se razloga najčešće koristi normalizirana vjerojatnost $p(N) = \frac{\langle k \rangle}{N}$.

Promotrimo distribuciju stupnja grafa $G_{n,p}$. Vjerojatnost da vrh v ima stupanj k jednaka je vjerojatnosti da je povezan s k različitih vrhova i nije povezan s $n-k-1$ preostalih vrhova. Primijetimo još i da je generiranje jednog brida neovisno od generiranja nekog drugog brida. Sada je jasno da je distribucija stupnja opisana binomnom distribucijom:

$$\mathrm{P}(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}.$$

Poznato je da binomna distribucija konvergira prema Poissonovoj kad $n \to \infty$ i kad je np fiksan. Stoga, za dovoljno veliki n i dovoljno mali p, Poissonovu distribuciju s parametrom $\lambda = np$ možemo koristiti kao aproksimaciju za binomnu distribuciju s parametrima n i p. Primijetimo da je u modelu $pn = \langle k \rangle$ konstantno, pa za dovoljno velike n koristimo aproksimaciju Poissonovom distribucijom:[10]

$$\mathrm{P}(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}.$$

Promotrimo sad neka osnovna svojstva ovog modela:

- Vjerojatnost da su proizvoljna dva susjeda nekog vrha v povezana jednaka je p. Stoga je koeficijent klasteriranja jednak $\langle C \rangle = \frac{np}{n} = p = \frac{\langle k \rangle}{n}$. Koeficijent se smanjuje širenjem grafa, pa zaključujemo da model nema ugrađen mehanizam grupiranja.

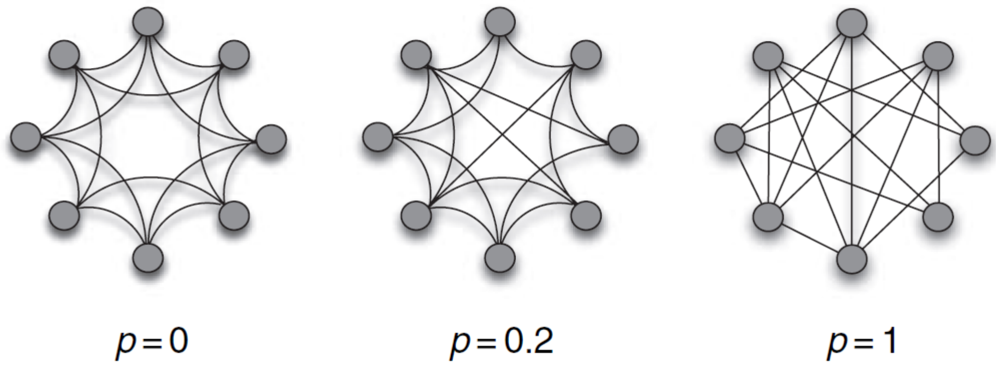
- Može se pokazati da prosječni najkraći put ima logaritamsku ovisnost o veličini grafa $\langle l \rangle \simeq \frac{\log n}{\log \langle k \rangle}$. Iz toga slijedi da u modelu nalazimo svojstvo malog svijeta.

- Distribucija stupnja može se opisati Poissonovom distribucijom koja spada među distribucije "lakog repa", pa zaključujemo da model opisuje homogenu mrežu.

Sljedeći model koji ćemo promatrati čine **generalizirani slučajni grafovi**. Model se također ponaša na slučajan način, ali s razlikom da možemo nametnuti proizvoljnu (moguću) distribuciju stupnja. Konstrukcija grafa teče tako da stupnjeve vrhova odredimo iz distribucije, pa u skladu s tim dodajemo bridove na slučajan način. Svojstva modela:

- Pokazuje se da koeficijent klasteriranja može poprimiti puno veće vrijednosti nego u Erdős–Renyi modelu, ali kako $n \to \infty$ koeficijent opet iščezava.
- Može se pokazati da prosječni najkraći put i dalje ima logaritamsku ovisnost o veličini grafa, pa je svojstvo malog svijeta i dalje prisutno. Dapače, pokazuje se da sama prisutnost nasumičnosti u izgradnji mreže inducira pojavu svojstva malog svijeta [7].
- Heterogenost mreže ovisi o distribuciji stupnja.

Mreže nastale od podataka u stvarnom svijetu često imaju vrlo jaki efekt grupiranja. Stoga se nameće potreba za traženjem modela u kojem će koeficijent klasteriranja to i prikazati. Jedan takav model je **Watts–Strogatz** model. Neka je n broj vrhova, $p \in [0, 1]$ i $\langle k \rangle = 2m$ željeni prosječni stupanj grafa. Graf konstruiramo tako da svaki od n vrhova simetrično povežemo s $2m$ najbližih susjeda. Tada sa svakim od m desnih susjeda ostajemo povezani s vjerojatnošću $1 - p$, odnosno nasumično biramo neki drugi vrh s vjerojatnošću p. Na taj smo način napravili prečace u grafu, dok smo istovremeno zadržali grupiranost s početka. Primijetimo da i kad $p \to 1$ graf nije statistički ekvivalentan slučajnom zbog činjenice da svaki vrh zadržava minimalno m bridova. Slika 1 prikazuje grafove iz Watts–Strogatz modela za $n = 8$ i $m = 2$.



Slika 1: Watts–Strogatz model za različite vrijednosti parametra p

Svojstva modela:

- Za $p = 0$ koeficijent klasteriranja vrlo je velik jer su susjedi svakog vrha međusobno jako dobro povezani. Za $p \to 1$ koeficijent klasteriranja iščezava zbog nasumičnosti.

- Jasno je da je za $p = 0$ prosječni najkraći put jako dug, $\langle l \rangle \sim n$, ali već za male vrijednosti p pojava prečaca drastično smanjuje put.

- Distribucija stupnja može se analitički izračunati i pokazuje se da je "lakog repa" te da mreža ima svojstva homogene mreže [3].

Napomena 4.2 *Zanimljivo je da postoji velik interval za izbor parametra p za koji će ova mreža imati svojstvo malog svijeta i velik koeficijent klasteriranja [15].*

Stanični model (eng. cellular model) prvi je među nabrojanim modelima koji podržava stvaranje komponenti, a kao i Watts–Strogatz model, potiče vrlo jako grupiranje. Model prima četiri parametra – broj vrhova, broj stanica, gustoću unutar stanica i gustoću izvan stanica. Konstruiranje grafa započinje podjelom vrhova po stanicama na nasumičan način. Nakon toga svaki vrh pripada točno jednoj stanici. Tada se svaka stanica gradi kao zasebna Erdős–Renyi mreža sa zadanom gustoćom unutar stanice. Nakon toga se stanice promatraju kao točke nove mreže. Potom se odlučuje koje će stanice biti povezane na isti način kao i u Erdős–Renyi mreži, u odnosu na zadanu gustoću izvan stanice. U slučaju da se dvije stanice povezuju, u svakoj stanici na nasumičan način odaberemo vrh koji će realizirati to povezivanje.

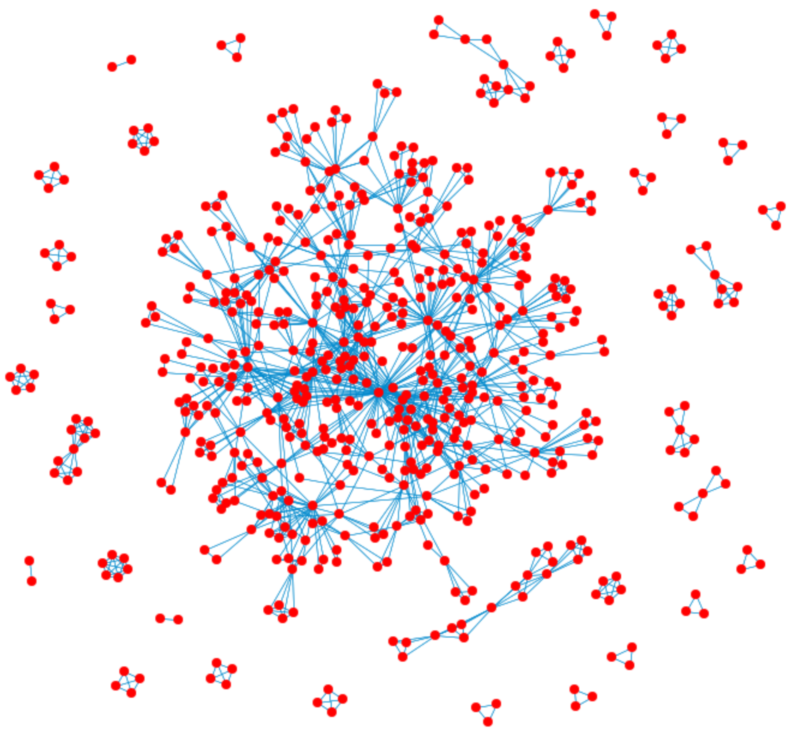
Svi dosad navedeni modeli bili su relativno strogo određeni. **Eksponencijalni slučajni grafovi** predstavljaju klasu modela koji se danas u svrhu modeliranja mreža najčešće koriste. Neka je n broj vrhova grafa i neka je X proizvoljna matrica susjedstva grafa s n vrhova. Promotrimo vjerojatnost da iz skupa svih matrica susjedstva grafova s n vrhova, u oznaci A_n, izaberemo baš matricu X. Tu vjerojatnost definiramo kao

$$\mathrm{P}(X) = \frac{\exp[\sum\limits_i \theta_i z_i(X)]}{\sum\limits_{Y \in A_n} \exp[\sum\limits_i \theta_i z_i(Y)]},$$

gdje su θ_i parametri modela, a $z_i(X)$ neke statističke opservacije. Prilagodljivost modela očituje se u opcijama pri određivanju statistika $z_i(X)$. Te veličine mogu predstavljati jednostavna svojstva grafa poput srednjeg stupnja $\langle l \rangle$, kao i komplicirana svojstva poput distribucije stupnja i/ili distribucije pripadnih atributa.

5 Analiza mreže radova objavljenih na DESIGN konferenciji

DESIGN mreža kao vrhove sadrži ključne riječi članaka objavljenih na DESIGN konferenciji znanosti o konstruiranju. Ključne riječi su prikupljene, usklađene od strane tima stručnjaka, te spremljene uz članak kojem pripadaju. Zatim je stvorena mreža koja povezuje ključne riječi ukoliko one pripadaju istom članku. Primijetimo da će mreža nastala na navedeni način biti reprezentirana neusmjerenim, multigrafom bez težina.



Slika 2: Mreža za 2014. godinu

5.1 Programsko rješenje

Predprocesiranje podataka i priprema za stvaranje mreže napravljena je u jeziku Python (`http://www.python.org`). Za modeliranje mreže i analizu nekih svojstava mreže korišten je specijalizirani alat ORA (`http://www.casos.cs.cmu.edu/projects/ora/index.php`) koji omogućava vizualizaciju i manipuliranje mrežama. Za analizu nekih svojstava mreže i statističku analizu korišten je R (`http://www.r-project.org/`). Za uspoređivanje s postojećim modelima mreže korišteni su je još i programski alat statnet (`http://statnet.org/`) i paket network(`http://cran.r-project.org/web/packages/network/network.pdf`) u R-u.

5.2 Dinamička analiza cijele mreže

Kako bismo vidjeli koliko se konferencije razlikuju iz godine u godinu, prvo ćemo promotriti mrežu na globalnoj razini.

U Tablici 1 dajemo neke opće informacije o mreži kroz godine. Kao što smo i vidjeli na Slici 2, ključne riječi se podijele u velik broj komponenti, dok povezanost, iako nije velika, ostaje solidna zbog velike centralne komponente. Gustoća grafa svake je godine prilično malena. Prosječna duljina najkraćeg puta svake je godine manja od $\log_2 n$. Ipak, iz toga ne možemo zaključiti da mreža ima svojstvo malog svijeta. Naime, za takav bi nam zaključak trebalo poznavanje asimptotskog ponašanja mreže, no ono nam nije dostupno. Naravno, valja naglasiti da smo u prethodnoj raspravi koristili malo izmijenjenu definiciju svojstva malog svijeta. Naime, ukoliko je graf razdvojen na nekoliko komponenata, postoje najmanje dva nepovezana čvora, pa je prosječna duljina najkraćeg puta u grafu beskonačna. Stoga za takav graf uopće nema smisla govoriti o navedenom svojstvu. U ovom je primjeru prosječna duljina najkraćeg puta računata samo za čvorove koji se povezani, pa ovdje svojstvo malog svijeta ima slično značenje kao i u prvotnom pokusu – ne može se doći od svakog do svakog čvora, ali ako se može, tada to možemo učiniti u malo koraka.

Tablica 1: Glavne vrijednosti po godinama

	2002	**2004**	**2006**	**2008**	**2010**	**2012**	**2014**
Broj vrhova	552	560	479	391	444	491	600
Gustoća	0.009	0.008	0.009	0.011	0.011	0.011	0.008
Broj komponenti	37	38	30	20	27	20	34
Povezanost	0.577	0.607	0.637	0.696	0.637	0.743	0.552
$\langle l \rangle$	4.352	4.177	4.150	4.604	3.981	3.962	4.319
Promjer najv. komp.	10	9	10	12	10	8	9
Koef. klasteriranja	0.843	0.836	0.838	0.804	0.812	0.809	0.866

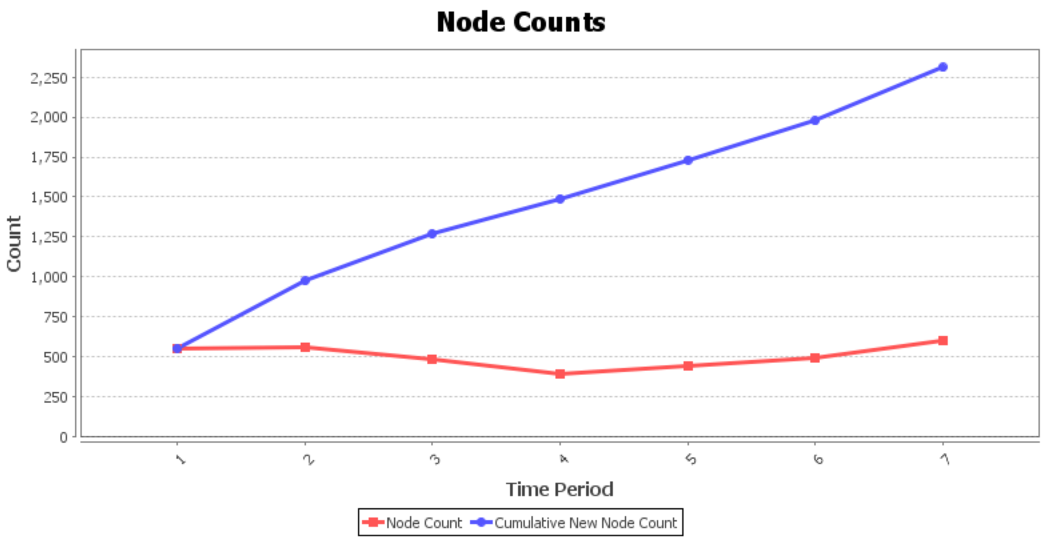
Slično, s obzirom na to da graf nije povezan, računali smo promjer najveće komponente grafa. Koeficijent klasteriranja je, unatoč maloj gustoći grafa, vrlo velik. No, to je bilo i očekivano razmislimo li kako je mreža nastala – od mnogo malih potpuno povezanih mreža. Dapače, pokazuje se da u svakoj od ovih mreža oko 70% vrhova ima koeficijent klasteriranja jednak 1 – njihovi su susjedi međusobno potpuno povezani.

U Tablici 2 vidimo analizu broja vrhova i bridova kroz mreže. Tako ukupno imamo 2311 različitih ključnih riječi, od čega se njih 551 pojavilo u barem dvije konferencije. U prosjeku se na konferenciji pojavi nešto više od 500 ključnih riječi, ali s relativno velikom standardnom devijacijom, čak 67,403. Zanimljivo je da se 33 ključnih riječi ponovilo na svim konferencijama, ali nijedna veza između dvije iste riječi nije se pojavila svake godine.

Tablica 2: Sličnost mreža

	Unija	**Djelomični presjek**	**Presjek**	**Sr. vr.**	**Std. dev.**
Broj vrhova	2311	551	33	502.429	67.403
Broj bridova	7925	316	0	1188.286	198.135

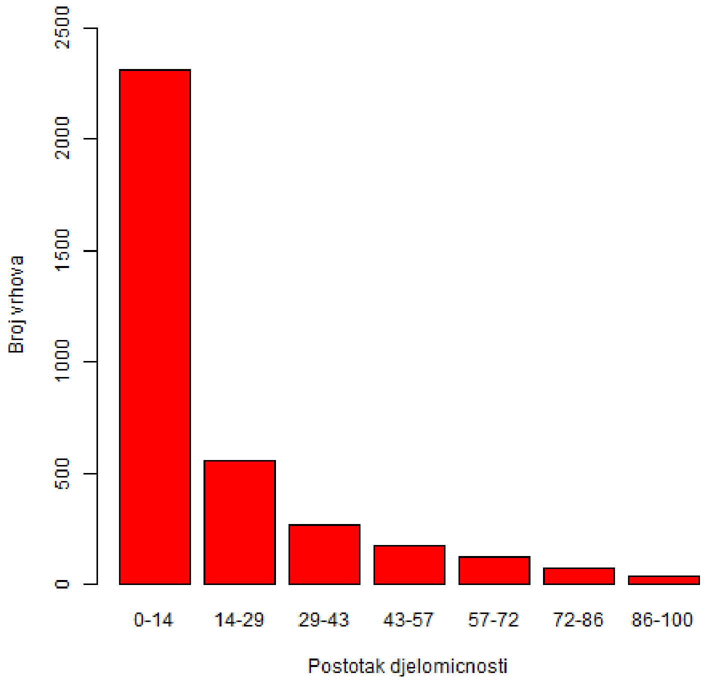
Na Slici 3 vidimo kako kumulativni broj vrhova svake godine raste, ali manje od broja vrhova koje mreža te godine sadrži. To je još jedan pokazatelj ponavljanja određenih ključnih riječi.



Slika 3: Broj vrhova i kumulativni broj vrhova po godinama

Djelomični presjek (eng. Lossy intersection) je presjek koji uzima u obzir određeni postotak broja mreža. Konkretno, u ovom slučaju to znači da za postotak djelomičnosti (eng. lossy percentage) do 14% gledamo sve vrhove koji se nalaze u barem jednoj mreži, dok za postotak djelomičnosti od 14 do 29% promatramo sve vrhove koji se nalaze u barem dvije različite mreže.

Na Slici 4 možemo vidjeti da se većina ključnih riječi pojavljuje na samo jednoj konferenciji.



Slika 4: Broj vrhova u odnosu na djelomičan presjek

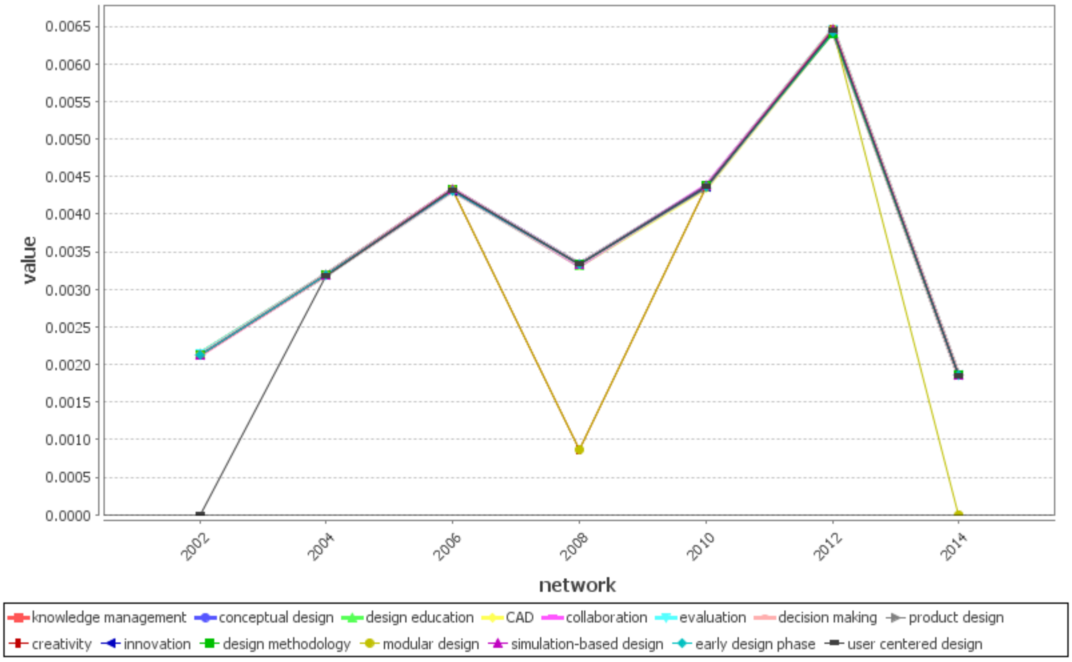
5.3 Dinamička analiza mreže po ključnim riječima

Dinamičkom analizom mreže utvrđujemo evoluciju istraživačkog područja temeljem radova objavljenih u sklopu DESIGN konferencije u razdoblju od 2002. do 2014. godine.

Analiza važnosti ključnih riječi za svako svojstvo prikazana je na 15 najvažnijih čvorova za dotično svojstvo.

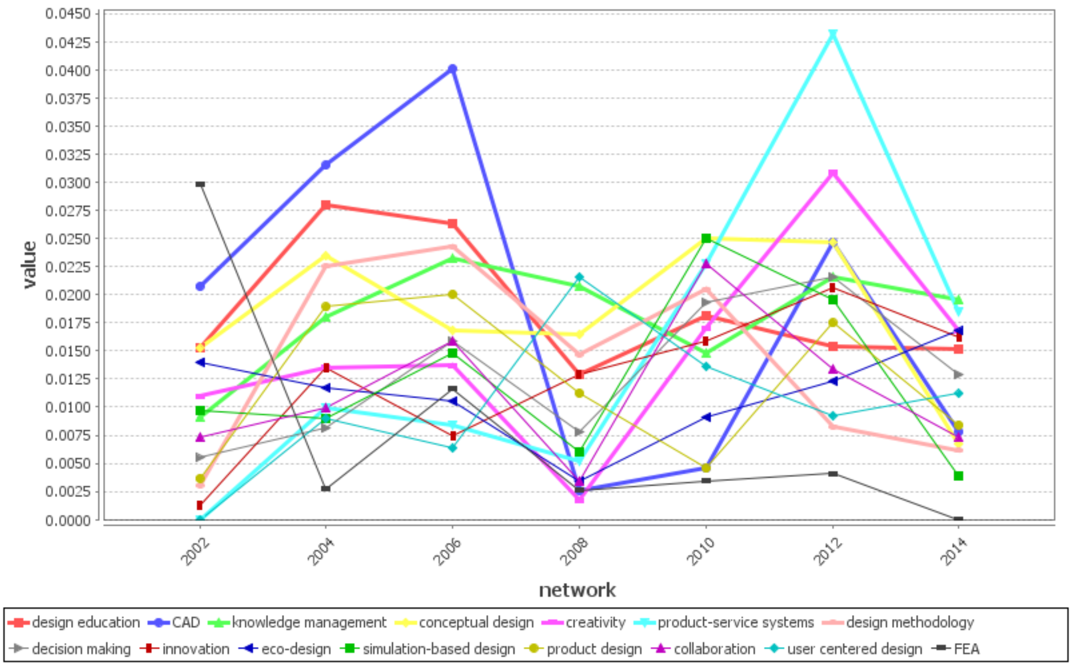
Na Slici 5 vidimo centralnost blizine ključnih riječi na konferencijama.

Slika 6 prikazuje ključne riječi s najviše veza prema drugim ključnim riječima. Na grafu se može vidjeti da je područje u razdoblju 2002-2014 doživjelo dva ciklusa konsolidacije. U prvom šestogodišnjem razdoblju (2002-2008) tradicionalne teme koje su bile važne za konstrukcijsku razradu komponenti proizvoda poput Finite



Slika 5: Centralnost blizine

Element Analysis (FEA) i Computer Aided Design (CAD) ili zahtjeva korisnika kroz User Centred Design, polako su gubile važnost ili su u potpunosti nestajale s liste tema kojima su se bavili istraživači u području, reflektirajući na taj način odmicanje područja od tradicionalne primijenjene mehanike, elementa strojeva i komercijalnih alata koji se koriste za konstrukcijsku razradu prema širem razvoju kompleksnih tehničkih sustava. U drugom šestogodišnjem razdoblju (2008-2014) došlo je do uspona tema koje su povezane s tom promjenom i promjenom paradigme razmišljanja o razvojnom procesu u širem kontekstu odnosa s okolišem i društvom u cjelini, te su teme koje dominiraju u tom razdoblju posvećene istraživanju sustava proizvod-usluga (PSS) u globalnom okruženju, kreativnosti i inovacija u razvoju tehničkih sustava te kriterijima konstrukcijske razrade važnima za održivost (Ecodesign). Posljednjih godina posebno se vidi utjecaj novih područja poput psihologije, kognitivnih znanosti, inženjerstva sustava, industrijskog dizajna, te menadžmenta koji su području dali nešto "mekši" predznak i odmak od tradicionalnog istraživanja u inženjerstvu. Tijekom cjelokupnog razdoblja promatranja, edukacija i metode za obrazovanje inženjera u ovom području, upravljanje inženjerskim znanjem, donošenje odluka (Decision making), konstruiranje temeljno na simulacijama te istraživanje ranih faza (Conceptual design) su konstantno prisutni, što pokazuje kako u tim temama istraživači još nisu pronašli zadovoljavajuće odgovore.

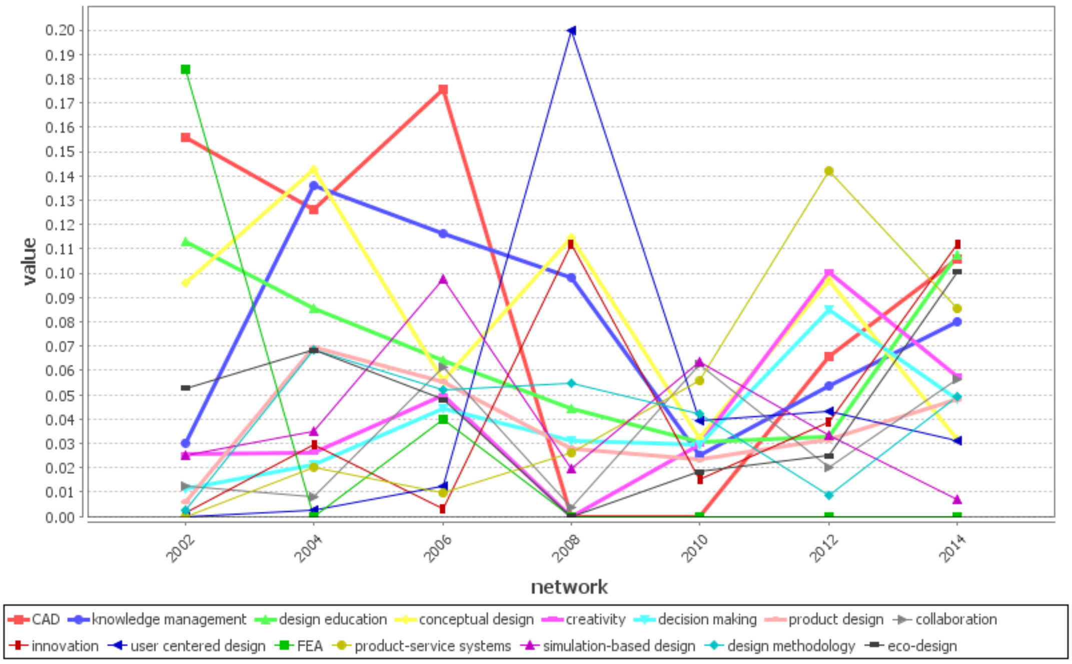


Slika 6: Centralnost stupnja

Vrlo male vrijednosti centralnosti bile su i očekivane s obzirom na to da je ova mreža ipak vrlo slabo povezana. Gotovo jednaka vrijednost centralnosti za svih 15 svojstava govori nam da sve te ključne riječi vjerojatno pripadaju istoj, najvećoj komponenti. Naime, udaljenost između dva nepovezana čvora definirana je kao $2n$, što je mnogo više od udaljenosti bilo koja dva čvora unutar komponente. Iz toga slijedi da ovu veličinu za čvor i dominantno određuje broj čvorova s kojima i nije povezan. To znači da svi navedeni čvorovi dolaze iz približno jednako velikih komponenti. Ali, kao što smo vidjeli na Slici 2, imamo samo jednu veliku komponentu.

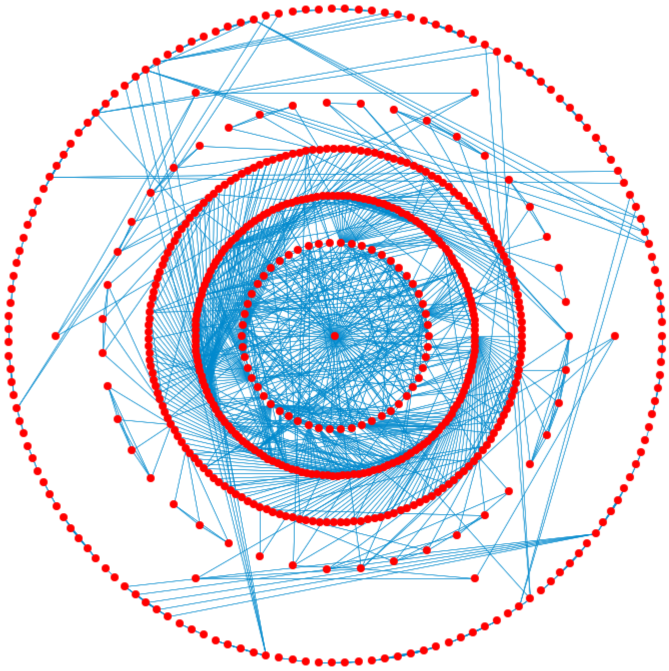
Ipak, zanimljivo je kako dvije ključne riječi iskaču u pojedinim godinama. To znači da je samo 13 ključnih riječi bilo unutar glavne komponentu kroz svih 7 konferencija.

Centralnost između vrhova ističe čvorove koji funkcioniraju kao mostovi, odnosno u ovom slučaju, ključne riječi koje povezuju različite grane znanosti. Stoga možemo reći da su ove ključne riječi “najviše interdisciplinarne”. Vidimo kako u zadnje vrijeme znanstvena područja najviše povezuju teme poput product-sevice systems, creativity, innovation i CAD.



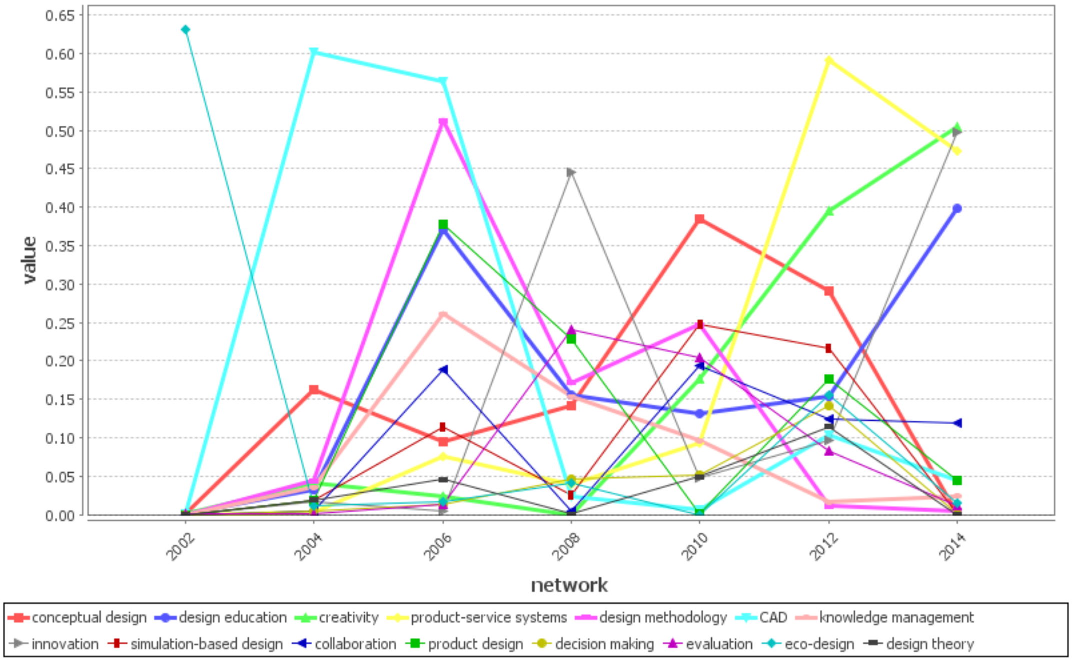
Slika 7: Centralnost između vrhova

Na Slici 8 vidimo kružni prikaz mreže iz 2014. godine. U centru se nalazi vrh s najvećim stupnjem centralnosti između vrhova, a na vanjskoj kružnici vrhovi s najmanjim stupnjem centralnosti. Ovakav prikaz naglašava važnost vrhova s visokom centralnosti između vrhova i činjenicu da služe kao mostovi u mreži.



Slika 8: Centralnost između vrhova na mreži DESIGN 2014.

Ideja centralnosti svojstvenog vektora je da su bitni oni vrhovi koji su dobro povezani s ostalim bitnim vrhovima. Ovaj naizgled cirkularan opis jasan je na primjeru internetskih stranica i algoritma Page Rank kojim tražilice poput Google-a određuju koje će stranice vratiti kao rezultate pretrage. Naime, navedeni algoritam stranicu smatra kvalitetnijom što je veći broj kvalitetnih stranica koje pokazuju na nju [9]. Na sličan način, ključne riječi koje se nalaze na ovom grafu predstavljaju na neki način najpopularnije ključne riječi na konferencijama. Kako su međusobno dobro povezane, znamo da se često pojavljuju zajedno u člancima, pa možemo reći da ti članci obrađuju popularne dijelove znanosti, odnosno te ključne riječi opisuju popularne dijelove znanosti. Na grafu se vidi značajan porast popularnosti ključnih riječi design education, innovation i creativity.



Slika 9: Centralnost svojstvenog vektora

5.4 Homogenost mreže

Neka nas svojstva zanimaju na razini statičke mreže. U tu ćemo svrhu promatrati mrežu radova iz 2014. godine.

Kako bismo klasificirali mrežu po homogenosti, promatrat ćemo funkciju distribucije stupnja i funkciju distribucije centralnosti između vrhova. U Tablici 3 prikazani su aritmetička sredina uzorka i uzoračka standardna devijacija. Njih ćemo po metodi maksimalne vjerodostojnosti koristiti kao procjenitelje za očekivanje i standardnu devijaciju.

Tablica 3: Svojstva distribucije stupnja i centralnosti između vrhova

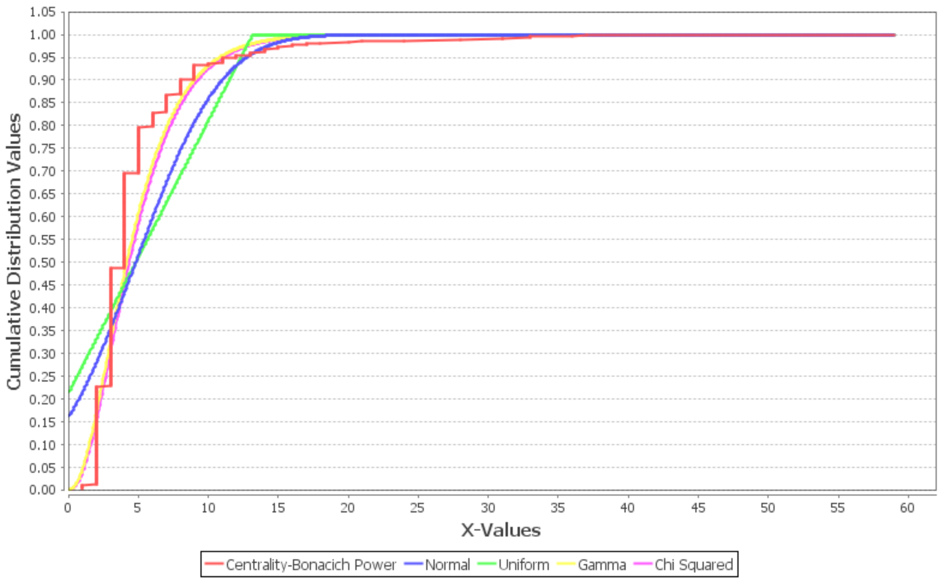
	Očekivanje	**Standardna devijacija**
Stupanj	4.82	4.82
Centralnost između vrhova	0.003	0.014

Funkciju distribucije stupnja prvo ćemo usporediti s nekima od najpopularnijih neprekidnih distribucija. U tu svrhu parametre smo računali pomoću metode maksimalne vjerodostojnosti.

Tablica 4: Procjena parametara metodom maksimalne vjerodostojnosti

Distribucija	1. parametar	2. parametar
Normalna	4.82	4.82
Uniformna	-3.53	13.17
Gamma	2.39	2.02
χ^2 **- distribucija**	5	/
Poissonova	4.82	/

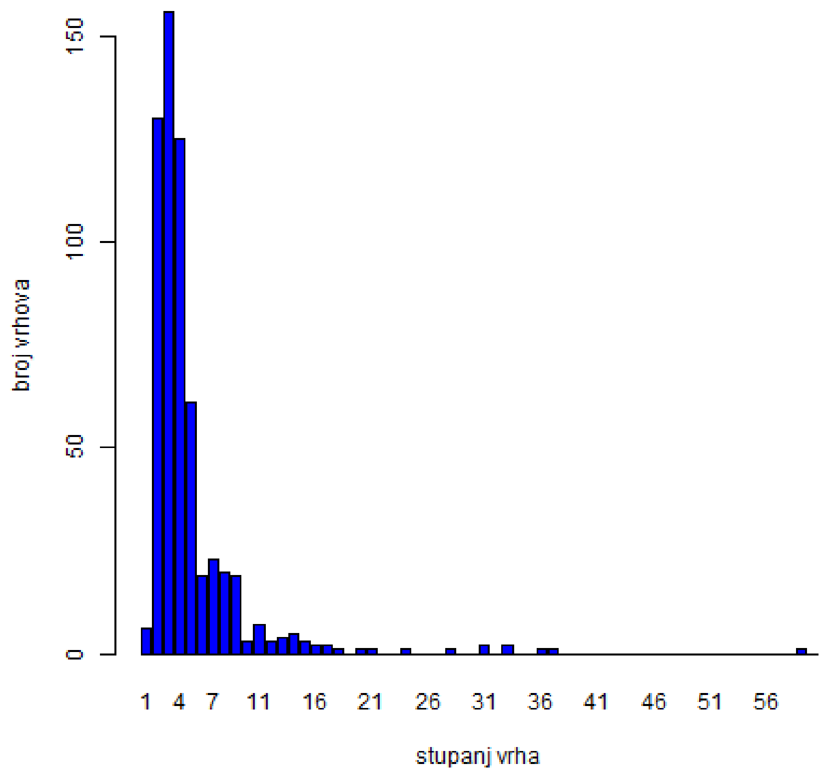
Slika 10 prikazuje usporedbu distribucije stupnja s neprekidnim distribucijama iz Tablice 4.



Slika 10: Usporedba distribucije stupnja s nekim neprekidnim distribucijama

Već gledajući sliku možemo naslutiti da se distribucija stupnja mreže ne može aproksimirati nijednom od navedenih distribucija. Tu su slutnju statistički testovi i potvrdili, svaki puta vrativši vrlo male p-vrijednosti ($< 2\times10^{-8}$), koje nam govore da ćemo hipotezu o pripadnosti tim distribucijama odbaciti za sve razumne razine značajnosti.

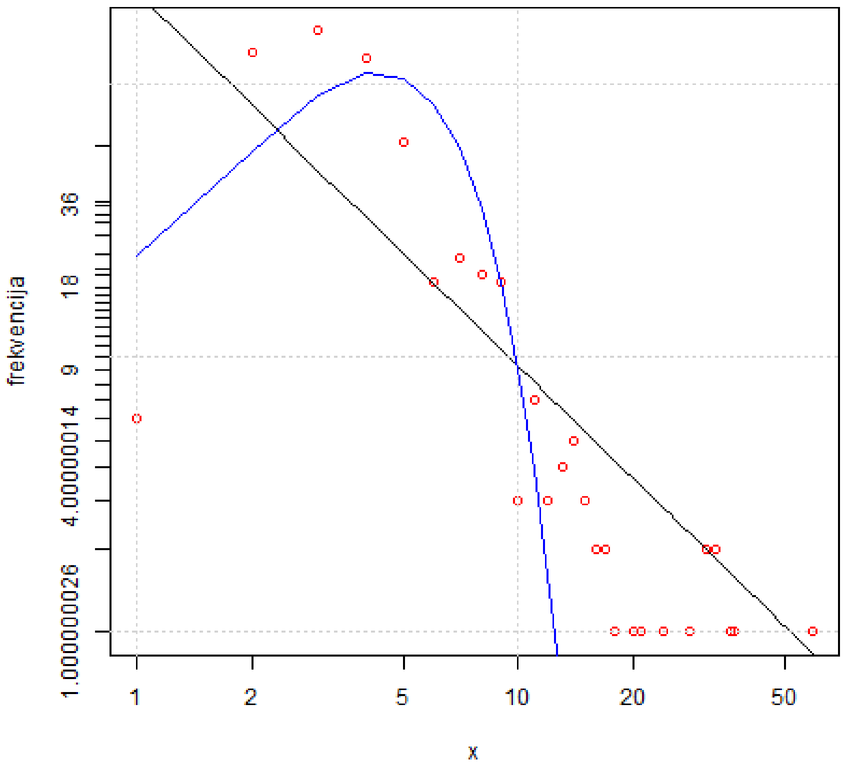
Promotrimo distribuciju stupnja na Slici 11.



Slika 11: Vrijednosti distribucije stupnja

Najčešća vrijednost stupnja u uzorku uvjerljivo je 3, dok aritmetička sredina uzorka nije toliko česta. Relativno velik broj realizacija vrlo daleko od aritmetičke sredine, potkrijepljen s razmakom u repu od ostalih distribucija na Slici 10, daje naslutiti da bi se moglo raditi o nagnutoj distribuciji teškog repa.

Slika 12 prikazuje usporedbu distribucije stupnja s distribucijom zakona potencija i Poissonovom distribucijom na dvostrukoj logaritamskoj skali. Crnom bojom označena je distribucija zakona potencija, a plavom Poissonova distribucija. Poissonova distribucija odabrana je kao predstavnik distribucija "lakog repa", a distribucija zakona potencija kao predstavnik distribucija "teškog repa". Naime, ove su dvije distribucije među najpopularnijima za opisivanje distribucije stupnja mreže i mogu se smatrati reprezentativnima za te dvije klase distribucija.



Slika 12: Usporedba distribucije stupnja s distribucijom zakona potencija i Poissonovom distribucijom

Pokazuje se da ni distribucija zakona potencija, kao ni Poissonova distribucija, nisu odgovarajuće distribucije, te se hipoteza o pripadnosti distribuciji u oba slučaja odbacuje zbog vrlo malih p-vrijednosti. S obzirom na to da testovi pripadnosti distribuciji nisu puno pokazali, potražit ćemo rješenje u koeficijentu heterogenosti. Tako računom za distribuciju stupnja dobivamo $\kappa = 9.64 \simeq 2\langle k \rangle$, a za distribuciju centralnosti između vrhova $\kappa = 0.068 \simeq 22\langle k \rangle$.

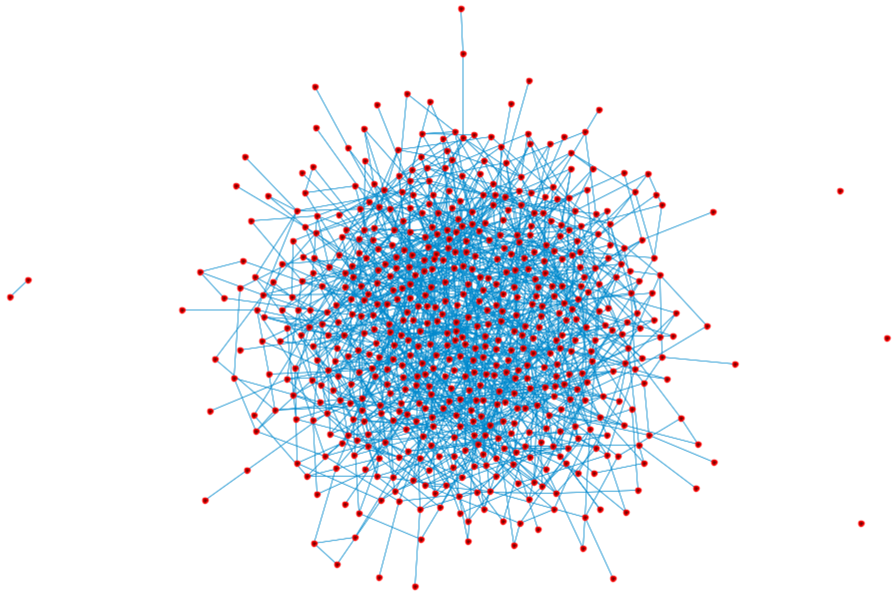
Vrijednost koeficijenta heterogenosti za distribuciju centralnosti između vrhova daje naslutiti da je tendencija mreže heterogenost, no zbog manjeg koeficijenta kod distribucije stupnja klasa se ne može odrediti sa sigurnošću, te ćemo u traženju modela promatrati i homogene i heterogene mreže.

6 Modeliranje mreže radova objavljenih na DESIGN konferencijama

Kako bismo modelirali mrežu, predstavit ćemo je kao jednostavan neusmjereni graf tako što ćemo promatrati samo egzistenciju veza između vrhova. S obzirom na vrlo velike promjene vrhova i bridova u svakom koraku mreže dinamički modeli mreža ne mogu opisati naše podatke. Stoga ćemo modeliranje komentirati na statičkoj mreži, i to onoj za 2014. godinu.

Iz teorije modeliranja mreža jasno je da Erdős–Renyi model ne može dobro opisati ovu mrežu. Naime, model je u potpunosti homogen i ima jako svojstvo malog svijeta, dok je našu mrežu bilo teško klasificirati po homogenosti, a kao što smo ranije i komentirali, svojstvo malog svijeta ne može imati u pravom smislu. Ipak, svojstvo u kojem vidimo najkritičniju razliku definitivno je koeficijent klasteriranja. Unatoč tome, potpunosti radi, nudimo usporedbu DESIGN mreže s najosnovnijim stohastičkim modelom.

Na Slici 13 je prikazana jedna Erdős–Renyi mreža s parametrom p izračunatim iz prosječnog stupnja vrha početne mreže. Mreža je razdvojena na komponente povezanosti. Iako je broj bridova doista ispao sličan broju bridova početne mreže (razlikuju se za dva brida), već je sa slike jasno da ova mreža ne može nikako opisati mrežu DESIGN 2014.



Slika 13: Erdős–Renyi mreža s prosječnim stupnjem vrha jednakom prosječnom stupnju vrha u mreži DESIGN 2014

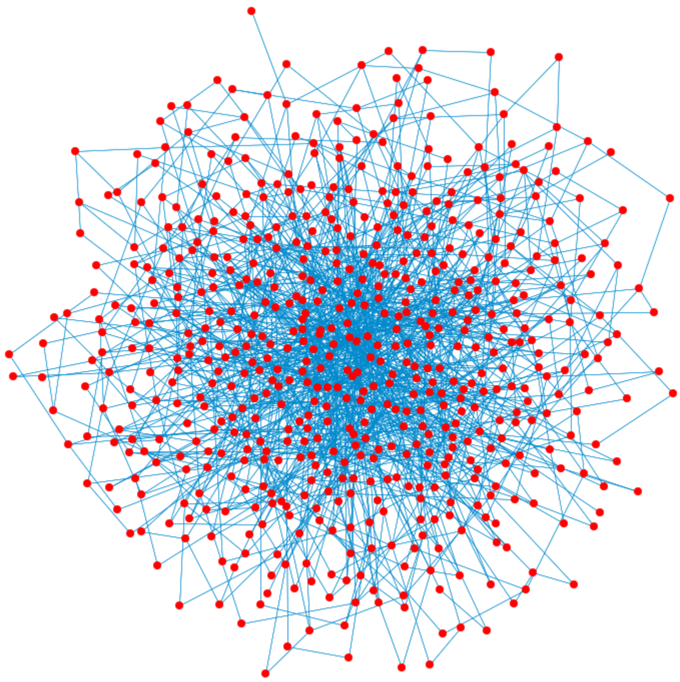
Da bi se to statistički i potvrdilo, napravljen je test jednakosti očekivanja nekih svojstava mreže u odnosu na svojstva modela. Dobiveni rezultati prikazani su u Tablici 5.

Tablica 5: Test jednakosti očekivanja s Erdős–Renyi modelom za neka svojstva mreže

Vrijednost	**DESIGN**	**Model**	**T-vrijednost**	**p.v.**
Centralnost između vrhova	0.003	0.001	2.692	0.043
Centralnost blizine	0.002	0.005	-128.482	$< 10^{-8}$
Inverzna centralnost blizine	0.14	0.02	36.46	3×10^{-7}

Model koji svakako valja uzeti u obzir jest model generaliziranih slučajnih grafova. Čini se da, kad bismo vrhovima dali onakve stupnjeve kakve imaju u originalnoj mreži, nije vjerojatno da dobiveni model jako odstupa. Promotrimo sliku

14 na kojoj se nalazi jedna mreža iz modela generaliziranih slučajnih grafova. Valja napomenuti da je ovo prikaz u kojem bi graf bio razdvojen na komponente povezanosti da ih ima. Jasno je da ovaj model ne uspjeva opisati mrežu DESIGN 2014.



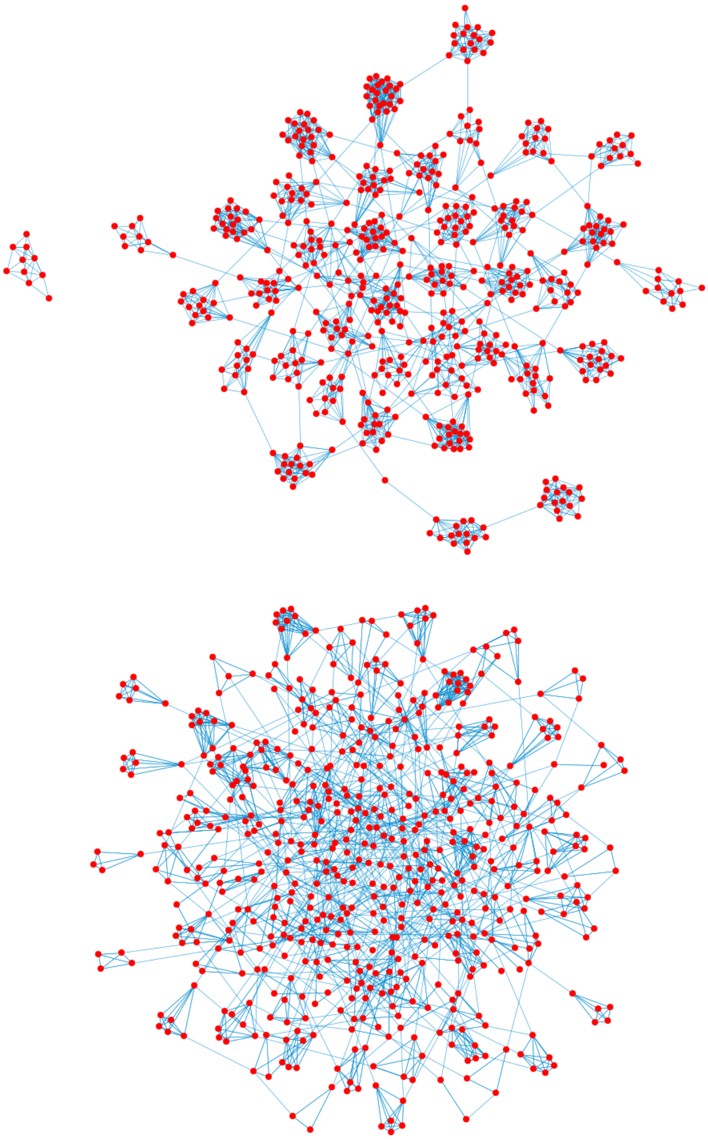
Slika 14: Model generaliziranih slučajnih grafova s jednakom distribucijom kao i u mreži DESIGN

DESIGN mreža sastoji se od velikog broja malih, vrlo povezanih mreža. Stoga se čini da ima smisla promatrati usporedbu sa staničnim modelom.

Slika 15 prikazuje dvije stanične mreže s različitim parametrima, ali s prosječnim stupnjem vrha jednakim prosječnom stupnju vrha u mreži DESIGN 2014. Te dvije mreže predstvaljaju dvije osnovne ideje o načinu na koji bi stanični model mogao aproksimirati originalni graf. Na prvoj slici nalazi se mreža izgrađena od tek nešto više stanica nego je originalni graf imao komponenti (40, odnosno 34). Ideja je bila staviti malu povezanost među stanicama u nadi da će mnogo komponenti ostati razdvojeno, ali da će se neke povezati u veću komponentu. Sada vidimo da, ne samo da takav parametar povezanosti izvan stanica ne postoji, već i da kad bi postojao, preostale komponente bile bi prevelike u odnosu na originalni graf, jer su sve stanice u početku imale podjednak broj vrhova (oko 15). Druga slika predstavlja mrežu izgrađenu od otprilike onoliko stanica koliko je bilo članaka na konferenciji (između 150 i 200). Ideja je da, uzmemo li velik broj komponenti s vrlo velikim stupnjem unutrašnje povezanosti, možemo dobiti strukturu sličnu člancima. To se pokazalo istinitim, no, stupanj vanjske povezanosti u tom će slučaju biti ili premalen da bi se konstruirala velika komponenta ili prevelik da bi velik broj stanica ipak ostao razdvojen po manjim komponentama. Ono što bismo htjeli je da neke stanice imaju vrlo malo veza prema drugim stanicama, a neke jako puno. Stoga bi se ovaj problem možda mogao zaobići tako da dopustimo zadavanje neke početne distribucije stupnja vanjske povezanosti. Naravno, još jedan potencijalni problem kod ovog pristupa stvara distribucija ključnih riječi po člancima u odnosu na distribuciju vrhova po stanicama. Sada je jasno da stanični model ne može opisati DESIGN mrežu, što se u Tablici 6 i vidi po rezultatima statističkih testova. Ipak, kad bismo htjeli definirati vlastiti model koji bi opisivao ovakve mreže, hibrid staničnog modela koji dopušta zadavanje početne distribucije povezanosti između stanica i/ili distribucije veličine stanica bio bi odlično mjesto za početak.

Tablica 6: Test jednakosti očekivanja sa staničnim modelom za neka svojstva mreže

Vrijednost	DESIGN	Model	T-vrijednost	p.v.
Centralnost između vrhova	0.003	0	5.290	0.003
Centralnost blizine	0.002	0	61.442	2×10^{-8}
Inverzna centralnost blizine	0.144	0	41.580	1.5×10^{-7}



Slika 15: Stanični model s jednakom srednjom distribucijom kao i u mreži DESIGN

7 Zaključak

Analiza mreže pokazala je kako su se glavne teme u znanosti o konstruiranju mijenjale kroz godine. Napravljena je analiza osnovnih svojstava mreže i analiza ključnih riječi kroz godine. Vidjeli smo da se posljednjih godina vidi značajan utjecaj novih područja poput psihologije, kognitivnih znanosti, inženjerskih sustava i drugih. Primijetili smo da mreža ima vrlo jaku sklonost grupiranju unatoč zapravo malenoj gustoći.

Proučavanjem statističkih svojstava mreže pokazali smo da mrežu ne možemo klasificirati po homogenosti, iako pokazuje blage tendencije heterogenosti. Iz te je činjenice jasno da je mreža prilično specifična. Naime, razlikuje se od slučajnih mreža koje uglavnom imaju vrlo jako svojstvo homogenosti, ali razlikuje se i od najčešćih mreža iz stvarnog života, koje većinom imaju izraženo svojstvo heterogenosti.

Usporedba mreže s Erdős–Renyi modelom, modelom generaliziranih slučajnih grafova i staničnim modelom pokazala je da postoji potreba za nekim drugim modelom koji bi mogao bolje opisati svojstva ove mreže. Predložen je hibrid staničnog modela koji bi uz određene izmjene mogao dobro opisivati ovu mrežu te se razrada tog modela ostavlja za daljnji razvoj.

8 Zahvale

Zahvaljujem mentoru prof. dr. sc. Mariu Štorgi i komentoru prof. dr. sc. Saši Singeru te prof. dr. sc. Sanji Singer na danim sugestijama i komentarima.

Također, zahvaljujem kolegama Anti Čabraji, Mariji Majdi Perišić, Dariu Bojanjcu i Tomislavu Drožđeku na vođenim raspravama za vrijeme pisanja rada i danim prijedlozima.

Literatura

[1] Figueredo, Ed Bullmore and Olaf Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems", Nature Reviews Neuroscience 2009, Vol. 10, 186–198.

[2] Fernando Vega–Redondo, "Complex Social Networks", Econometric Society Monograph Series, Cambridge University Press, 2007.

[3] Barrat, A. and Weigt, M., "On the properties of small-world network models", Eur. Phys. J. B 13 (2000), 547–560.

[4] Pavlos Antoniou and Andreas Pitsillides, "Understanding Complex Systems: A Communication Networks Perspective", Technical Report TR–07–01, Department of Computer Science, University of Cyprus, February 2007.

[5] Washington Y. Ochieng, Mohammed A. Quddus, Robert B. Noland, "Map-matching in complex urban road networks", Brazilian Journal of Cartography, (2003) 55 (2), 1–14.

[6] Stanley Milgram, "The Small World Problem", Psychology Today, 1967, Vol. 2, 60–67.

[7] Alain Barrat, Marc Barthelemy, Alessandro Vespignani, "Dynamical processes on complex networks", Cambridge University Press, 2008.

[8] A.-L. Barabasi and R. Albert, "Emergence of scaling in random networks", 1999, Science 286, 509–512.

[9] Sergey Brin and Lawrence Page, "The Anatomy of a Large-Scale Hypertextual Web Search Engine", Stanford University, `http://infolab.stanford.edu/~backrub/google.html`.

[10] Simons, Gordon; Johnson, N. L. "On the Convergence of Binomial to Poisson Distributions", The Annals of Mathematical Statistics 42 (1971), no. 5, 1735–1736.

[11] M. E. J. Newman, "The structure and function of complex networks"

[12] Bin Xiong,Zhongyi Zheng, "Graph theory", Mathematical Olympiad series, Vol 3.

[13] P. Erdős and A. Renyi, "On random graphs", Publ. Math. 6, 1959, 290–297.

[14] Leo Spizzirri, "Justification and Application of Eigenvector Centrality", `https://www.math.washington.edu/~morrow/336_11/papers/leo.pdf`.

[15] D. J. Watts, S. H. Strogatz, "Collective dynamics of 'small-world' networks", 1998, Nature 393 (6684): 440–442.

[16] E. N. Gilbert, "Random graphs", 1959, Ann. Math. Stat. 30, 1141–1144.

9 Sažetak

U ovom je radu analizom kroz godine promatran razvoj znanosti o konstruiranju, na temelju konferencija održanih od 2002. do 2014. godine. Svaka je konferencija, pomoću ključnih riječi u člancima koji su na njoj predstavljeni, modelirana kao jedna statička mreža. Zatim su te mreže promatrane kroz godine i objedinjene u jednu dinamičku mrežu. Ispitana su osnovna svojstva mreže na razini vrhova te izdvojene ključne riječi koje na temelju različitih poznatih mjera (centralnosti) predstavljaju najvažnija područja znanosti. Ta su područja promatrana kroz godine te je komentiran razvoj znanosti.

Za potrebe rada u poglavlju Reprezentacija i analiza mreža izneseni su osnovni pojmovi teorije grafova i neke veličine bitne za analizu mreža. Zatim su promotrena osnovna svojstva mreža. Predstavljeni su osnovni modeli mreža i njihove karakteristike.

Osim analize na razini vrhova, napravljena je usporedba mreža na globalnoj razini i promatrane promjene raznih karakterističnih veličina poput broja komponenti, koeficijenta klasteriranja i gustoće. Na mreži koja predstavlja zadnju konferenciju, onu iz 2014. godine, analizirana je homogenost mreže, proučavane su distribucije stupnja i centralnosti između vrhova te je napravljena usporedba s osnovnim poznatim vjerojatnosnim distribucijama. Na istoj mreži napravljena je i usporedba s nekim modelima mreža i predložen hibrid staničnog modela čiji je razvoj ostavljen daljnjem istraživanju.

10 Summary

This work observed the evolution of Design science, analysing it through the years, based on conferences that took place from 2002 through 2014. Each conference was modelled as a single static network, using the keywords of the articles presented on a forementioned conference. Afterwards, we observed the obtained networks through the years and put them together into a single dynamic network. Basic network properties (on a vertex level) were examined, and based on various known measures (e.g. centrality), keywords representing the most prominent branches of science were singled out. These branches were then observed year by year, annotating the science development.

In the section 3, basic terms of Graph Theory were laid out, along with some values vital for network analysis. Basic network properties were also observed later in the work. Some network models, along with their important characteristics were represented.

Besides the vertex-level analysis, we also compared networks on a global scale, observing changes of various characteristic values, such as number of components, clustering coefficient and density. On a network representing the 2014 conference, we analysed the network homogeneity, studied the degree distribution and betweenness distribution, and afterwards compared the obtained results to some well known probability distributions. We also compared the said network to various network models and recommended a cellular model hybrid, whose development was left to be further researched in subsequent works.