



NORGE

(12) **PATENT**

(19) NO

(11) **321340**

(13) **B1**

(51) Int Cl.

*H04L 12/24 (2006.01)*

### Patentstyret

---

(21)	Søknadsnr	20035852	(86)	Int.inng.dag og søknadsnr	
(22)	Inng.dag	2003.12.30	(85)	Videreføringsdag	
(24)	Løpedag	2003.12.30	(30)	Prioritet	Ingen
(41)	Alm.tilgj	2005.07.01			
(45)	Meddelt	2006.05.02			
(73)	Innehaver	Telenor ASA , Snarøyveien 30, 1331 FORNEBU, NO			
(72)	Oppfinner	Geoffrey S Canright, Waldemar Thranesgate 40B, 0171 OSLO, NO Kenth Engø-Monsen, St Olavs gate 3, 3126 TØNSBERG, NO Åsmund Welzien, Dælenenggata 21C, 0567 OSLO, NO			
(74)	Fullmektig	Oslo Patentkontor AS , Postboks 7007 Majorstua, 0306 OSLO, NO			

---

(54)	Benevnelse	<b>Fremgangsmåte for å administrere nettverk ved analyse av konnektivitet</b>			
(56)	Anførte publikasjoner	D1: Gkantsidis, C.; Mihail, M.; Zegura, E, "Spectral analysis of Internet topologies", IEEE INFOCOM 2003 D2: WO-A2 03/058868 D3: Bonacich, P; Lloyd, P.: "Eigenvector-like Measures of Centrality for Asymmetric Relations", University of California at Los Angeles (udatert) D4: WOA1 03/036914 D5: US-A1 2003/0147400 D6: Davis, M; Wilson, C.S.: "Visualizing the cooperation networks of elite institutions in India", 8th International Conference on Scientometrics and Informetrics. Proceedings - ISSI-2001 D7: Tuva Stang, Faimeh Pourbayat, Park Burgess, Geoffrey Canright, Kenth Engø, Åsmund Welzien, "Archipelago: A network security analysis tool", in proceedings of the 17th Annual Large Installation Systems Administration Conferenc (LISA 2003)			
(57)	Sammen drag				

Det er beskrevet en fremgangsmåte for å bestemme et nettverks evne til å spre informasjon eller fysisk trafikk. Nevnte nettverk omfatter et antall nettverksnoder som er sammenknyttet av linker. Fremgangsmåten omfatter å kartlegge topologien av et nettverk, beregne en verdi for linkstyrke mellom nodene, beregne en egenvektor sentralitetsindeks for alle noder, basert på nevnte linkstyrkeverdier, identifisere noder som er lokale maksima av Egenvektor sentralitetsindeksen som senternoder, gruppere nodene i regioner som omgir hver identifisert senternode, tildele en rolle til hver node fra sin posisjon i en region, som senternoder, regionmedlemsnoder, grensenoder, bronoder, dinglenoder, og måle nettverkets mottagelighet for spredning, basert på antall regioner, deres størrelse, og hvordan de er koblet.

## **Oppfinnelsens område**

Foreliggende oppfinnelse dreier seg om et sett  
5 fremgangsmåter for administrering av nettverk (både logiske  
og fysiske nettverk), innen et antall områder. Nærmere  
bestemt beskriver foreliggende oppfinnelse en fremgangsmåte  
for analysering av et nettverk, hvor nettverket består av  
ethvert antall nettverksnoder sammenknyttet av linker.

10

## **Bakgrunn for oppfinnelsen**

Nesten alle sosiale eller fysiske strukturer hvor  
individuelle enheter er linket sammen av en type forhold  
15 kan analyseres fra et nettverksperspektiv, om det er  
sosiale grupper, flyruter, eller grupper av datamaskiner.  
Nettverk er interessante objekter. De har en god del  
struktur, samtidig som de er enkle. De består, i sin  
enkleste form, bare av noder sammenknyttet ved linker. Den  
20 abstrakte ide av et nettverk, eller grafe - uttrykkene  
brukes om hverandre - er også meget nyttig ved modellering  
av strukturer observert i den virkelige verden. Eksempler  
omfatter: Sosiale nettverk, kommunikasjonsnett,  
"verdensveven" (the World Wide Web), metaboliske og  
25 genetiske nettverk i biologiske system, matnettverk,  
sykdomsnettverk, og kraftforsyningsnett. Kort sagt er et  
nettverk en enkel, ikke-triviell abstrakt struktur,  
fascinerende i seg selv, og også meget relevant for mange  
grener av vitenskap og teknologi.

30

Innen området telekommunikasjon har teorier vedrørende  
nettverksstyring og nettverksstrukturer vært etablert i  
lang tid. Det er av avgjørende viktighet å forstå et  
nettverk. Effektiviteten av drift og vedlikehold av et  
35 slikt telekommunikasjonsnett vil hovedsakelig avhenge av  
kjennskap til det aktuelle nettverk. Det er viktig både med  
henblikk på midlere tid mellom svikt, så vel som med

henblikk på spredning av skade, slik som virus, ormer eller lignende.

For datakommunikasjonsnett er situasjonen mye den samme.

- 5 Lignende betraktninger er relevant for drift av elektriske strømforsyningsnett, særlig med henblikk på sikkerhet. Innen planlegging og drift av elektriske nett er det viktig å ha et robust nettverk, og slik for eksempel unngå situasjoner hvor en stor del av en populasjon utsettes for
- 10 strømutfall. Analyse av konnektiviteten av et nettverk er viktig for betraktninger rundt robusthet.

Systemadministrasjon involverer uvergelig styring av et nettverk, som er sammensatt av flere typer linker.

- 15 Eksempler omfatter: De fysiske linker mellom maskiner, de logiske linker mellom brukere og filer, og de sosiale linker mellom brukere. Et viktig aspekt ved systemadministrasjon er å sikre den frie flyt av nødvendig informasjon over nettverket, mens man samtidig hindrer flyt
- 20 av skadelig eller ødeleggende informasjon over det samme nettverk.

Strukturen av nettverket spiller en avgjørende rolle i implementeringen av disse to viktige, og delvis

- 25 motstridende, mål for systemadministrasjon. Begge mål involverer spredning av informasjon over linker i nettverket; derfor er begge problemer sterkt følsomme for nettverksstrukturen. På grunn av denne avhengigheten kan forståelsen av nettverksstrukturen være en verdifull
- 30 komponent i effektiv systemadministrasjon.

Videre er det selvfølgelig nettverk som er både sosiale og teknologiske. Eksempler inkluderer telefonnett; "peer-to-peer" nettverk [10] som ligger på Internett; og det

35 kombinerte nettverk av datamaskiner, filer og brukere som er den daglige oppgave for enhver systemadministrator.

Her synes sikkerhet igjen å være et åpenbart bruksområde for disse ideer: Man ønsker å identifisere noder som bør gis høyest prioritet for beskyttelse mot virus, for eksempel.

5

Studier av nettverk har mottatt en stor del oppmerksomhet i det siste tiåret eller så. De fleste mål for nettverksstruktur som har blitt studert til nå [8], er i form av "helgrafe" egenskaper, det vil si skalare mål eller fordelinger som gjelder grafen som helhet, og som beregnes ved bruk av gjennomsnitt. Eksempler på slik helgrafe egenskaper inkluderer nodegradfordelingen, diameteren eller midlere banelengde, klyngekoefisienter, og notasjonen av "små verdener", som er basert på de to foregående.

15

Helgrafe egenskaper er viktige og nyttige; imidlertid kan de ikke gi et fullstendig svar på spørsmålet "Hvordan kan vi forstå strukturen av et nettverk?"

20 Det eksisterer mange eksempler hvor kjennskap til nettverk, som tar en mer abstrakt form enn i telekom, datakom eller elektriske nett, er av viktighet. For eksempel er det på området epidemiologi viktig å ha en forståelse av sosiale nettverk og hvordan disse nettverk forenkler spredning av sykdommer. Innen informasjonsdistribusjon er det av viktighet å kjenne mekanismene som styrer spredningen av informasjon innen en populasjon, om det er på et lokalt eller globalt nivå.

30 Når man ser på inter-humane forhold eller sosiale nettverk, ser man på linkene mellom individene i stedet for deres kategorier eller hva som karakteriserer dem. Et sosialt nettverk er derfor enhver gruppe av personer hvor individene har en form for relasjon til hver andre. Personer med en høy grad av sosial påvirkning i sosiale nettverk kalles ofte opinionsledere. De har påvirkningskraft enten på grunn av deres ekspertise eller på grunn av deres sosiale posisjon. I alle fall vil denne påvirkningskraft ofte vise

seg ved å gi opinionsledere et stort antall sosiale kontakter; de er *linket* med et stort antall mennesker. Dette er selvfølgelig logisk; å ha sosial påvirkningskraft betyr at du har evnen til å nå et stort antall mennesker.

5

Nytten av denne ideen for *sosiale nettverk* synes klar [4]. Det er åpenbart også av interesse å identifisere samfunn i et målt sosialt nettverk. Et eksempel med et noe annet anstrøk er nettverk av seksuelle kontakter. Også her kan disse ideer være ganske nyttige, i arbeid rettet mot å  
10 begrense spredning av seksuelt overførte sykdommer: Kanskje man skulle fokusere på de to komplementære mål av (i) å hindre infeksjon av sentralnodene i hvert samfunn, og (ii) hindre overføringen av sykdom over bronodene.

15

Av disse årsaker fortjener nettverk seriøse undersøkelser. Et nettverk er et av de enkleste abstraksjoner av struktur som kan studeres; likevel er forståelse av strukturen i et nettverk en ikke-triviell oppgave.

20

### **Kjent teknikk**

På det vitenskapelige området nettverksanalyse er det flere måter for å måle sentraliteten av nettverksnoder. Et av  
25 disse mål kalles *egenvektor sentralitet*. Egenvektor sentralitet (EVC for "Eigenvector Centrality") ble definert tidlig i syttiårene av Bonacich [2]. Grunnideen bak EVC er, at det ikke bare er hvor mange mennesker du kjenner, men også hvor viktig (sentrale) de er, som bestemmer hvor  
30 viktig (sentral) du er. Dette er derfor faktisk en rekursiv definisjon; min viktighet (sentralitet) avhenger av mine naboers' - som i sin tur avhenger av min. Poenget med en slik rekursiv definisjon er å tillate oss å identifisere de nav (hubs) som virkelig er innflytelsesrike fra  
35 perspektivet til hele nettverket. Ellers ville en definisjon som telte viktighet bare etter hvor mange naboer du har, løpe risikoen for å nominere sentere av *isolerte klynger* som nettverksnav. Med henblikk på sosiale nettverk

er disse sentre innflytelsesrike bare i en begrenset forstand, siden deres innflytelse ikke strekker seg ut over deres umiddelbare naboer.

5 Arbeidet til Kleinberg [7], mens det gjelder nettverk med rettede linker, gir noen nyttige perspektiver. Kleinberg betraktet en rettet grafe, som definerte to distinkte typer roller for nodene på grafen, og ga en måte å beregne indisier som kvantifiserer i hvilken grad hver node spiller  
10 de to typer av roller. Det vil si at hver node i en rettet grafe kan tildeles Autoritetspoeng og Navpoeng. Det er viktig å legge merke at disse poeng kan være basert bare på topologien av grafen - uavhengig av ethvert spørsmål om innhold eller betydning, eller av enhver "egenskap" til  
15 nodene.

Navnene til disse to rolletyper formidler deres betydning. Noder med høy autoritet er noder som er viktige fordi de pekes på av viktige noder - faktisk av noder med høye  
20 navpoeng. Og de siste oppnår deres høye navpoeng ved å peke mot gode autoritetsnoder. Kort sagt: Nav peker, og autoriteter pekes på. Disse ideer kan være meget nyttige for å analysere strukturen av WWW: Autoriteter er sannsynligvis gode endepunkter for et informasjonssøk, mens  
25 nav er nyttige for å lede søket til autoritetene. Det synes klart at tilsvarende roller oppstår i sosiale nettverk: Noen ganger vet man hvem som har den nødvendige informasjon (autoriteten), andre ganger må man spørre et godt nav.

30 Kleinbergs arbeid gir oss *indisier* for hver node i nettverket. Disse indisier gir oss nyttig informasjon om rollen(e) noden spiller i nettverket. Slik informasjon er ganske distinkt fra helgrafe informasjon; og enda er den fremdeles utledet, i det minste slik som opprinnelig  
35 presentert, fra den rent topologiske struktur av grafen.

Et annet aspekt ved grafen, som igjen er distinkt fra helgrafe egenskaper, er samfunnsstrukturen til grafen. I

den samme artikkel foreslår Kleinberg en måte å finne slike samfunn i grafer slik som Web-grafen. De matematiske verktøy som brukes er grunnleggende de samme som de som brukes til å finne Nav/Autoritetspoeng - hvilket blant  
5 annet betyr at dekomponeringen av grafen i samfunn også var fullstendig basert på strukturen til grafen, uten å påkalle noen kjennskap eller egenskaper til nodene eller linkene. Videre kan det legges merke til at dekomponering av en grafe i sub-samfunn gir ny informasjon om rollene som  
10 spilles av noder; de kan være medlemmer av samfunn X; de kan ligge i intet samfunn; de kan være "ledere" av deres samfunn i noen forstand, eller de kan ligge på "kanten"; og de kan spille en viktig rolle for å linke multiple samfunn.

15 Mange andre arbeider har adressert det samme problem med hvordan man finner "naturlige" samfunn i en rettet grafe slik som Webben. I motsetning til dette har Girvan og Newman [5] sett på dette spørsmål for ikke-rettede grafer. Deres basale fremgangsmåte er å definere samfunn ved først  
20 å finne deres "grenser" - nærmere bestemt ved å finne linker med høy "mellomhet", som om de fjernes bryter grafen opp i sub-samfunn.

Fra Stang [13] er det kjent en fremgangsmåte for å  
25 analysere et nettverk omfattende å kartlegge topologien til nettverket, beregne en verdi for linkstyrke mellom noder i nettverket, beregne en Eigenvektor sentralitesindeks for alle nodene basert på linkstyrkeverdiene, identifisering av noder som er lokale maksima i Eigenvektor  
30 sentralitesindeksen som senternoder, gruppering av nodene i regioner rundt hver senternode, tildeling av roller til hver node fra sin posisjon i en region. Imidlertid viser ikke artikkelen hvordan fremgangsmåten kan implementeres i praksis, og går ikke nærmere inn på de viktige spørsmål  
35 hvordan linkstyrke mellom noder skal beregnes, eller hvordan nodene grupperes i regioner.

### Sammenfatning av oppfinnelsen

Det er en hensikt med foreliggende oppfinnelse å tilveiebringe en fremgangsmåte for nettverksanalyse, som kan benyttes enten på fysiske nettverk, eller på logiske nettverk som eksisterer som overliggende nettverk på toppen av det fysiske nettverket. Det viktige felles aspekt er identifikasjonen av linker (fysiske eller logiske), over hvilke informasjonen kan flyte.

10

En annen hensikt medforeliggende oppfinnelse er å tilveiebringe et "naturlig" middel - det vil si et som bare er basert på grafetopologien - for å dekomponere en ikke-rettet grafe inn i samfunn. Et sett roller for nodene i grafen vil bli definert, slik at hver node blir tildelt en, og bare en, rolle. I motsetning til fremgangsmåten til Kleinberg, er det for det foreliggende bruksområde ønskelig at rollene er binære (Ja/Nei) egenskaper av nodene - og eksklusive i tillegg.

20

Tidligere kjent teknikk [13, 3] har i større detalj vist hvordan man benytter analysen presentert her på datamaskiner i nettverk med mange brukere. Foreliggende oppfinnelse tilveiebringer en naturlig måte å dekomponere et nettverk i vel sammenknyttede klynger, og av å tildele meningsfulle roller i informasjonsflyt til hver node i nettverket.

Disse hensikter oppnås i en fremgangsmåte for nettverksanalyse som beskrevet i det vedføyde patentkrav 1. Nærmere bestemt tilveiebringer oppfinnelsen en fremgangsmåte for analyse av et nettverks evne til å spre informasjon eller fysisk trafikk, idet nettverket omfatter et antall nettverksnoder sammenknyttet av linker, idet fremgangsmåten omfatter trinnene:

35

- kartlegging av topologien av et nettverk,
- beregning av en verdi for linkstyrke mellom nodene,



- beregning av en Egenvektor Sentralitetsindeks for alle noder basert på nevnte linkstyrkeverdier,
- identifisering av noder som er lokale maksima av Egenvektor Sentralitetsindeksen som senternoder,
- 5 • gruppering av nodene i regioner som omgir hver identifisert senternode,
- tildeling av en rolle til hver node fra dens posisjon i en region, som senternoder, regionmedlemsnoder, grensenoder, bronoder, dinglenoder,
- 10 • måling av nettverkets mottagelighet for spredning, basert på antall regioner, deres størrelse, og hvordan de er sammenknyttet,
- tildeling av rollen som regionmedlemsnoder i en gitt region til alle noder for hvilke den bratteste
- 15 oppstigningslinkbane i topologikartet terminerer unikt i senternoden for denne region.

Fordelaktige utførelser av oppfinnelsen fremgår av de etterfølgende uselvstendige krav.

20

#### **Kortfattet beskrivelse av figurene**

For å gjøre oppfinnelsen mer lett forståelig, vil oppfinnelsen nå bli diskutert i detalj med henvisning til de vedføyde figurer, hvor:

25

Figur 1 er et skjematisk diagram som viser en Bronode (venstre) og Bronode og Dinglere (høyre).

30 Figur 2 viser en enkel grafe med to regioner.

Figur 3 viser den samme grafe som i Figur 2, men med regionene definert ved bruk av en annen regel. EVC-verdier for nodene er også vist.

35

Figurene 4, 5 og 6 viser de resulterende grafer av MANA-prosjektet [4] ved bruk av tre forskjellige mål for linkstyrke.

Figur 7 er et flytdiagram som illustrerer fremgangsmåten benyttet for å beregne Egenvektor sentralitetindeksen.

## 5 Detaljert beskrivelse av oppfinnelsen

Nyttige og interessante bruksområder for ideer for nettverksanalyse beskrives av foreliggende oppfinnelse. Det eneste kravet er at linkene i nettverket er ikke rettede. Med ikke-rettede linker mener vi linker som ikke peker i en bestemt retning. På World Wide Web kan en web-side peke til en annen, men denne side må ikke nødvendigvis peke tilbake. I dette tilfelle ville sidene være forbundet av en rettet link. Hvis begge sider var hyperlinket til hverandre, men en link i hver retning, ville disse linker kollapse til en ikke-rettet link. Foreliggende oppfinnelse behandler alle nettverk som bestående av ikke-rettede linker.

Ideen fulgt av foreliggende oppfinnelse er at "veltilknyttethet" kan betraktes som en høydefunksjon over det diskrete rom (grafene). Hvis høydefunksjonen til foreliggende oppfinnelse er glatt nok, kan det benyttes ideer passende for glatte overflater over et kontinuerlig rom. Det vil si at foreliggende oppfinnelse vil benytte et topologisk bilde for å definere regioner i et nettverk. Regioner vil korrespondere til "fjell", med senteret i hver region som den korresponderende fjelltopp. Grenser mellom regioner vil deretter defineres som de punkter som ikke er unikt assosiert med en fjellregion.

De definerte roller er: "Leder" av et samfunn (region); medlem av et samfunn; og to typer roller for noder i "grensesettet", det vil si noder som ikke tilhører noe samfunn.

Fremgangsmåten er grovt regnet motsvarende den til Girvan og Newman [5]. Foreliggende oppfinnelse begynner ikke med "kantene", men med "sentrene" i samfunnene. Fra dette

startpunkt arbeider man seg "utover" for å finne medlemmene, og til slutt grensenodene. Det presenterte sett av roller er fullstendig og konsistent, i den forstand at definisjonene tillater en unik og entydig assosiasjon av en enkelt rolle til hver node i grafen.

### **Utførelse av foreliggende oppfinnelse**

Mennesker som kommuniserer med hverandre danner et sosialt nettverk, hvor linkene er basert på deres kommunikasjon. Disse linker kan adskilles ifølge det type medium som benyttes, det være seg telefoni, ansikt-til-ansikt kommunikasjon, eller post. Slik kan det sosiale nettverk beskrives som *multipleks*: Det er et nettverk hvor nodene er relatert til hverandre ved forskjellige typer linker. Selv om de sosiale forhold som linker forskjellige personer sammen kan eksistere uavhengig av den type medium som benyttes, spiller typen medium en viktig rolle for å definere linkene, etter som hvert medium er en distinkt kanal for informasjonsflyt. Forskjellige kommunikasjonsmedier er i denne forstand analoge til språk. For eksempel må en person som ønsker å nå mange noder i nettverket være i stand til å kommunisere over multiple typer av medier - han må snakke de andre noderes foretrukne "språk". Denne ide med linker differensiert av media gjelder for de fleste typer nettverk: Sykdom kan for eksempel spre seg gjennom et antall forskjellige bærere av infeksjoner, og linker i transportnett kan bestå av mange forskjellige transportmidler, for eksempel biler, fly eller tog.

### **Linkstyrke og EVC-mål**

Styrken av linkene i denne type multipleks nettverk kan bestemmes på forskjellige måter. Her nevner vi fire:

- 1) Man kan simpelthen angi om en link (av enhver type) eksisterer eller ikke. Numerisk tildeler man 0 til "ikke link" og 1 til "noen link".
- 5 2) Man kan telle *antallet forskjellige medier* som knytter sammen ethvert par av noder, det vil si antallet forskjellige medier som har noen flyt av substanser eller informasjon mellom enhver to noder i nettverket.
- 10 3) Man kan måle den *totale flyt* mellom enhver to noder i nettverket. For å gjøre dette må man konvertere dataene som er tilgjengelige til et felles mål. Dette mål gir derfor nettomengen av flyt (for eksempel minutter eller ord for kommunikasjonsmedia) mellom enhver to noder i nettverket.
- 15 4) Et fjerde alternativ er å bestemme styrken av linkene gjennom en *blanding* av 2) og 3). Det vil si å telle hvert medium [som i 2)], men vektet [som i 3)] av flytfraksjonen for dette medium, som et gitt par bruker.
- 20 Den tradisjonelle måte å bestemme linkstyrke på er indikert som nummer 1). Metode 3) er også kjent. Metodene 2) og 4) er nye og innovative metoder for bestemmelse av linkstyrke.
- 25 Egenvektor sentralitets- (EVC) indeksen defineres matematisk som hovedegenvektoren av en matrise. Den enkleste og mest vanlige metode for å finne hovedegenvektoren av en matrise er den "opphøyde metode" [14]. Denne metode involverer gjentatt multiplikasjon på en vektor av vektorer med matrisen. Multiplikasjon på vektvektoren av matrisen er ekvivalent med det som kalles "vektforplantning": Det omfordeler et sett vektorer ifølge en regel. gjentatt omfordeling av vektene (med total normalisering av den totale vekt) gir en stabil fordeling, som er den dominante eller hovedegenvektoren. Dette er poengene som brukes som sentralitetsindeks av foreliggende oppfinnelse.
- 30
- 35

For klarhets skyld illustrerer vi bruk av den Opphøyde Metode, i Fig. 7. Her, med bruk av sammenhengene forklart tidligere, starter prosessen og en startvektor  $w_0$  blir valgt (s401). Ved hver gjentakelse blir en ny vekt  $w_{new}$  beregnet (s403) ved omfordeling av vektene ifølge virkningen av matriseoperatoren. Denne nye vekt blir deretter normalisert (s405). En konvergenstest blir deretter utført (s407). Hvis vekten har konvergert, avsluttes prosessen. Ellers blir en ny vekt beregnet og prosessen gjentatt inntil vekten konvergerer.

For analyse av multiplekse sosiale nettverk, har EVC-målet blitt generalisert til å inkorporere tre andre mål for linkstyrke (2-4), som nevnt ovenfor. Modifikasjonen av den generelle EVC-ide, som benyttet i de nye metoder 2) og 4), er som følger: En node er sentral hvis den har mange naboer med høy sentralitet - og bruker mange forskjellige typer av media. I det etterfølgende blir det beskrevet hvordan denne generelle ide implementeres for hver av de fire tilnærmelser til linkstyrke diskutert ovenfor:

1) Den tradisjonelle tilnærming, hvor naboskapsmatrisen  $\underline{A}$  er dannet av bare 0'ere og 1'ere, kan brukes med multipleks nettverk; men den er totalt ufølsom for antallet media som brukes av hvert par av noder.

2) Her erstatter vi simpelthen matrisen  $\underline{A}$ , hvis poster alle er enten 0 eller 1, med matrisen  $\underline{A}_{color}$ , definert som følger: Posten ( $\underline{A}_{color}$ ) er lik antallet "farger" (distinkte media) som knytter sammen nodene  $i$  og  $j$ .

3) Her er 1'erne i den tradisjonelle  $\underline{A}$  matrisen erstattet med et positivt reelt tall, gitt det totale volum av flyt (summert over alle media, og målt i en felles enhet av mål) over et gitt tidsintervall. Det vil si:

$$(\underline{A}_{volume})_{ij} = \sum_c V_{c,ij}$$

hvor  $c$  er en indeks som strekker seg over "farger" (media), og  $V_{c,ij}$  er det totale kommunikasjonsvolum i medium  $c$  mellom nodene  $i$  og  $j$ .

4) Endelig foreslår foreliggende oppfinnelse en blanding av tilnærmelser 2) og 3), for å gi vekt til både flytvolum og til eksistensen av multiple media. Derfor, for hvert medium  $c$  og nodepar  $ij$ , gir vi "poeng" som er fraksjonen (bidratt av paret  $ij$ ) av den totale kommunikasjon som bruker medium  $c$  i nettverket. La  $V_{T,c}$  betegne det totale volum (over hele nettverket) av kommunikasjon som bruker medium  $c$ . Da kan vårt "blandede" mål av linkstyrke skrives som  $(A_{mixed})_{ij} = \sum_c (V_{c,ij} / V_{T,c})$ .

10

Fremgangsmåten ifølge foreliggende oppfinnelse konverterer flytdata til en naboskapsmatrise, ved bruk av de fire metoder beskrevet ovenfor, for å gi hver matrisepost et linkstyrkemål. Deretter beregner den hovedegenvektoren av den resulterende modifiserte naboskapsmatrise. Dette tillater oss å tildele et indeks (et positivt tall) til alle nodene i nettverket, gitt deres sentralitet ifølge et av våre fire mål. Nodene med de høyeste sentralitetsverdier betraktes som de mest sentrale noder i nettverket. Dette tillater fremgangsmåten å frembringe et topologisk kart av nettverksstrukturen, det vil si en grafisk visualisering av nettverket som viser de mest sentrale noder som lokale "topper".

## 25 **Roller i nettverk**

Det endelige mål med foreliggende oppfinnelse er å tildele en naturlig og unik rolle til hver node i nettverket, basert på topologien av grafen alene. Som bemerket ovenfor, fant Kleinberg to slike roller for rettede grafer: Nav og Autoriteter. Nav er naturlig gode til å peke mot gode Autoriteter; og Autoriteter er naturlig gode til å bli pekt på av gode Nav. Man kan allerede se fra disse enkle grammatikalske utsagn at distinksjonen mellom Nav og Autoriteter forsvinner når buene i grafen blir ikke-rettet (slik at "peker på" = "blir pekt på"). Matematikken gir det samme resultat: For det ikke-rettede tilfelle er naboskapsmatrisen symmetrisk,  $A=A_T$ , og slik blir matrisene som definerer Nav og Autoriteter de samme.

35

Kort sagt, for ikke-rettede grafer, vil de to typer av roller kollapse til en. Denne ene rolle (mer presist, en indeks som kvantifiserer i hvilken grad noden spiller rollen) er egenvektor sentralitet.

Navoperatoren  $AA^T$  og Autoritetsoperatoren  $A^T A_{\text{simply}}$  blir  $A^2$ , hvis hovedegenvektor er den samme som den til  $A$ .

Det er derfor funnet at to av rollene identifisert i Kleinbergs arbeid med rettede grafer blir en enkelt (type av) rolle for en ikke-rettet grafe. Denne rolletype kalles veltilknyttethet i de etterfølgende avsnitt, eller egenvektor sentralitet. Det blir videre søkt etter distinksjoner blant nodene for en ikke-rettet grafe - med andre ord multiple distinkte roller, til hvilke enhver gitt node kan tildeles. Disse roller vil bli definert i neste avsnitt. Egenvektor sentralitet (EVC) vil være høydefunksjonen, og derfor startpunktet.

20

### **Definisjoner av rollene**

Forskjellen mellom "rolletype" og "rolle" må klarlegges. Indisier eller "poeng" med reelle verdier kan assosieres med hver node: Nav og Autoritespoeng for det rettede tilfelle, og EVC-poeng for det ikke-rettede tilfelle. Dette er rolletyper; faktisk er det rettferdig å si at alle tre poeng representerer en type sentralitet. Alle noder har en grad av sentralitet; og å "være sentral" er bestemt en type rolle. Imidlertid, med rolle menes i dette dokument en binær (ja/nei) distinksjon brukt på hver node, slik at hver node mottar et enkelt Ja og derfor tildeles en unik og entydig rolle. Sentralitet (en rolletype) vil gi en glatt høydefunksjon over grafen, og tillater bruk av topografiske kriterier for å tildele en (JA eller Nei) rolle til hver node.

35

### Sentre

For enkelhet og lesbarhet beholdes bildet av fjell, daler, saler osv. for høydefunksjonen. Hvert fjell kan defineres ved sin topp. Toppen er et lokalt maksimum av høydefunksjonen. Den første rolle er da fjelltoppen.

**Senter:** *Enhver node som er et lokalt maksimum av egenvektorsentraliteten er et senter.*

10

### Regioner

Hver fjelltopp definerer et fjell. Derfor er antallet regioner i grafen lik antallet sentre. (I det etterfølgende blir store bokstaver utelatt, unntatt når roller blir definert; betydningen skulle fremgå av sammenhengen.) Regioner er vanligvis sammensatt av mer enn en node; derfor kan ikke rollen for en node være en region, men i stedet et regionmedlem.

20

**Regionmedlem:** *Hver node som unikt kan assosieres med et enkelt senter, ifølge en enhetlighetsregel, er et medlem av dette Senterers region, og derfor et Regionmedlem.*

25 Det gjenstår å definere "enhetlighetsregelen". Ifølge foreliggende oppfinnelse er det gitt to mulige valg for "enhetlighetsregelen".

30 **Regel 1 (avstand).** *En node er assosiert med Senter C hvis den er nærmere (i antall korteste bane hopp) til C enn til ethvert annet senter  $C_0$ .*

35 **Regel 2 (brattest oppstigning).** *For hver node i vil en bane med brattest oppstigning som starter i i terminere i et (eller flere) Sentre. Hvis den terminerer i et enkelt Senter er node i assosiert med dette Senter.*



Disse regler er simpelthen diskret-domene versjonen av prosessen med å assosiere en del av domenet (basisrom) med hver fjelltopp - og slik definere hvert fjell. Man må være nøyaktig her for å bryte definisjonen av region i to deler:

5 Selve definisjonen, som viser til en regel men ikke spesifiserer den; og regelen. Dette gjøres fordi mer enn en regel er mulig for det diskrete tilfelle; og regiondefinisjonen er angitt på en måte som fanger inn "fjellideen", men lar regelen være uspesifisert.

10

Begge regler angitt ovenfor tilfredsstillende det intuitivt fornuftige kriterium at et senters nære naboer (generelt) bør tilhøre sin region. (Det er uansett antallet og tilknyttetheten til et senters naboer som gir dette senter sin høye EVC.) Begge regler er også enkle å implementere på

15 en enkel iterativ måte - startende med sentrene, og arbeidende utover fra dem, "fargende" noder ifølge regionene (sentrene) de tilhører. Bratteste-oppstigningsregelen er imidlertid den regel som er mest trofast mot det

20 topologiske bilde.

### **Grenser - mellom regioner**

På en kontinuerlig topografisk overflate er det punkter som ligger mellom fjell, og tilhører intet unikt fjell. Det kan

25 være at analoge punkter eksisterer for det diskrete tilfelle også.

Noder som ikke kan assosieres med noe fjell tildeles

30 Grensesettet.

**Grensenoder:** *Enhver node for hvilken enhetlighetsregelen for Regionmedlemsskap gir mer enn et svar er en Grensenode.*

35 Intuitivt tenker man på grensesnoder som "koblende regioner". Og likevel vil litt mer tenkning avdekke at ikke alle grensenoder er like i dette henseende. Noen grensenoder spiller faktisk en viktig rolle med å koble to

eller flere regioner: De ligger på baner som kobler de respektive sentre (derfor regioner). Se venstre bilde i Figur 1. Andre noder kan fjernes uten noe tap i graden av forbindelse mellom regionene. Se høyre bilde i Figur 1.

5 Derfor er det naturlig å definere to distinkte roller til settet av grensenoder.

**Bronoder:** *En Bronode som ligger på i det minste en ikke-selv-gjenfindende bane som kobler to Sentre er en Bronode.*

10

**Dingler:** *Enhver Grensenode som ikke er en Bronode er en Dingler.*

Dinglere kan selvfølgelig injisere ny informasjon inn i nettvirket; men de spiller ingen betydelig rolle i transporten av informasjon mellom regionene.

Endelig er det ønskelig å peke ut en klasse av linker som spiller en viktig rolle i kobling av regioner. Årsaken for å gjøre slik her er at grensesettet for bratteste-oppstignings-regelen generelt er meget lite eller null. I dette tilfelle er det fremdeles nyttig å understreke de nettverkselementer som kobler regionene. derfor blir det definert:

25

**Brolinker:** *Enhver link hvis endepunkter ligger i to distinkte Regioner er en Brolink.*

Brolinker vil opptre for begge regionregler ovenfor. Man kan forestille seg regler for å definere regioner som gir "fete" grenser. For eksempel kan man assosiere noder med sentre ifølge:

**Regel 1' (distanse med kutt).** *En node assosieres med Senter C hvis den er nærmere (i antall hopp) til C enn til ethvert annet Senter  $C_0$ , og hvis dens avstand fra C ikke er større enn h hopp.*

"Fete" grenser oppstår for en slik regel siden det kan være mange noder som er fjernere enn h hopp fra ethvert senter. Generelt oppstår "fete" grenser hvis man velger en regel innrettet til å unngå "sammengroing" av regioner fra deres respektive sentre. Avstand til hvilken vekst tillates kan da måles i hopp (som i Regel 1'), eller i minskning i EVC.

Grenser ifølge Regel 1 er "tynne": Essensielt en node bred. Grenser ifølge Regel 2 er til og med tynnere: Generelt er de 0 noder brede, siden det er sjelden at en node vil ha to eller flere bratteste oppstignings baner, førende til forskjellige lokale maksima.

### Matematikken

15

De matematiske problemer som løses av foreliggende oppfinnelse løses med fokus på "glatte" funksjoner over et diskret rom.

20 Anta at domenerommet er kontinuerlig. Da er harmoniske funksjoner de glatteste funksjoner som er tilgjengelig. Disse funksjoner er løsninger av Laplace's ligning,

$$(1) \nabla^2 \phi = 0$$

25

For et gitt rom oppnår man forskjellige løsninger for (1) fra forskjellige grensebetingelser på  $\phi$ .

Man vil umiddelbart identifisere noen problemer med kontinuum-bildet. et problem er at det ikke er noen maksima, eller minima, vekk fra grensen. derfor kan ikke det topografiske bilde ifølge foreliggende oppfinnelse virke med slike glatte funksjoner: Hver fjelltopp vil ligge på grensen. Videre beskriver foreliggende oppfinnelse en naturlig måte å definere regioner på. Her betyr "naturlig" ledet så mye som mulig av topologien til grafene. derfor er det uønsket å måtte tildele verdier for funksjonen  $\phi$  ved

grensen - det vil foretrekkes at topologien løser dette problem.

5 Man kan selvfølgelig løse dette siste problem ved å sette  $\varphi$ =konstant, for eksempel null, ved grensen. Det vil si at grensen bare gis en nominell grenseverdi. Dette er "naturlig" nok; imidlertid får man da at  $\varphi$ =konstant over hele rommet, på grunn av midlingsegenskapene til Laplace's ligning.

10

Den diskrete versjon av Laplace's ligning er

$$(2) \quad L\varphi=0$$

15

hvor  $L=K-A$  er Laplace-matrisen,  $K=\text{Diag}(k_1, k_2, \dots)$  er en diagonalmatrise hvis  $i$ te post er nodegraden  $k_i$ , hvor  $k_i$  er antallet koblede naboer av node  $i$ , og  $A$  er naboskapsmatrisen, med  $A_{ij}=1$  hvis det er en link fra  $i$  til  $j$ , og 0 ellers.

20

Det er lett å se at midlingsegenskapene holder her også: Løsninger til (2) følger

$$(3) \quad \phi_i = \frac{1}{k_i} \sum_{j=m(i)} \phi_j$$

25

Her betyr "nn" "nær nabo". Den diskrete Laplace-ligning tilbyr derfor "mest glatt" funksjoner for det diskrete tilfelle; men den har alle problemer sett for kontinuerlige harmoniske funksjoner, pluss et til. Tilleggsproblemet stammer fra det avgjørende faktum at spesifikasjonen av grensen av et diskret rom ikke er unik - faktisk er det ingen naturlig måte å definere en slik grense. Man kan selvfølgelig ta den kanskje minst tilfeldige forutsetning at ingen av punktene er grensepunkter - alle har deres høyde bestemt av grafestrukturen - men da får man tilbake konstanten  $\phi_i$ =konstant.

### Eigenvektor sentralitet

Følgende diskusjonen fra uttrykk (3). En liten endring i bildet som gis av (3) løser alle dens problemer samtidig.

5 Den lille endringen er som følger: Det spørres etter en høydefunksjon som adlyder, i stedet for den midlende egenskap (3), det følgende:

$$(4) \quad \phi_i = \frac{1}{\lambda} \sum_{j=nn(i)} \phi_j$$

10

Det vil si at i stedet for å ta det strikte middel over alle naboer, deler man nabosummen med en konstant  $\lambda$ , som er den samme for alle noder. Ligningen kan da skrives som

$$15 \quad (5) \quad A\phi = \lambda\phi$$

Hvor  $A$  igjen er naboskapsmatrisen. Nå har vi en egenverdiligning, og høydefunksjonen  $\phi$  er en egenvektor av naboskapsmatrisen. Foreliggende oppfinnelse ønsker faktisk egenvektoren som er den stabile iterative løsning av (4), fordi høyde antas å bety "veltilknyttethet". Det vil si at (4) koder den idé at node  $i$  sin veltilknyttethet er bestemt, til innen en skalakonstant  $\lambda$ , av det til alle  $i$ 's naboer. Gjentakelse av dette krav, fra ethvert startpunkt, vil gi hovedegenvektoren for naboskapsmatrisen. Denne egenvektor gir den stabile selv-konsistente løsning av (4); den har også den egenskap at den er positiv halv-endelig, siden  $A$  er.

30 Med denne ene modifikasjon er problemene sett ovenfor med Laplaces ligning (diskrete eller ellers) ikke lenger til stede. EVC kan ha lokale maksima vekk fra grensen. Siden den måler veltilknyttethet, tenderer faktisk lokale maksima av EVC å ligge godt borte fra alle noder som man kan  
35 fristes å kalle "grensenoder". Videre er det ikke behov for å definere en grense for det diskrete tilfelle: Alle noder

kan ha EVC-verdier bestemt av ligning (4), med ingen verdier inngitt som "grensebetingelser".

Spesifikt er bidragene her:

5

1) De to nye modifiserte former for naboskapsmatrisen, som gir to nye mål for sentralitet som tillater nettverkssentre å plukkes ut.

10

2) Definisjonen og metoden for å identifisere nettverksregioner.

3) Definisjonene og metodene for å tildele diskrete nettverksroller til hver node i nettverket.

15

4) Anvendelse av de nye mål for sentralitet, regioner og roller på et vidt utvalg bruksområder.

### Eksempler

20

I det etterfølgende gis det eksempler på utførelser av foreliggende oppfinnelse så vel som sammenligninger mellom de to regler for å definere regioner.

25

Figurene 4, 5 og 6 viser resultatene av MANA forskningsprosjektet som presenteres i [4]. Grafene representerer et lite sosialt nettverk, en arbeidsgruppe med 11 personer. Med bruk av den presenterte fremgangsmåtes forskjellige mål for linkstyrke, ble EVC-baserte sentralitetsindisier laget for nettverket. De topografiske visualiseringer viser sentraliteten av nodene som forskjeller i høyde. I Figur 4 er linkstyrke målt basert på antallet forskjellige media som brukes av hver node (metode 2). Figur 5 viser grafen når linkstyrken er basert på netto mengde flyt mellom nodene (metode 3). Endelig viser Figur 6 grafen som er basert på en blanding av metodene ovenfor for å bestemme linkstyrke, det vil si både antall media som brukes og netto mengde flyt (metode 4).

30

Figur 2 viser en enkel grafe med to sentre. Grensen består av tre noder. En node (node 11) er en bronode som klart spiller en essensiell rolle i å koble de to regioner, de andre to er dingle. 5

Bruk av regel 2 på den samme grafe gir oss Figur 3. Her kan man se at hele grensen har blitt "slukt" av det dominante senter (node 9). Den heller perifere rolle til noder 10 og 12 - tidligere klassifisert som dingle - reflekteres nå i deres avstand (2 hopp) fra deres senter (og selvfølgelig i deres lave EVC). 10

Sammenligning av disse to figurer bekrefter forventningene om forskjellene mellom de to regler: et grensesett, med eller uten dingle, er typisk til stede med Regel 1, men fraværende med regel 2. 15

For å illustrere anvendelsen av disse ideer, antar vi at nodene i Figur 2 og 3 er brukere i et datanettverk, mens linkene er effektive forbindelser mellom brukere som tillater informasjonsflyt. Her brukes uttrykket "effektive" forbindelser, fordi linkene ikke behøver være direkte: de kan formidles av filer til hvilke begge brukere har lese- og skrivetilgang. 20  
3. Man kan konkludere umiddelbart fra analysen at brukersystemet er naturlig sammensatt av to hovedgrupper. Videre er node 9 mest sentral for den gule gruppen, mens node 13 er mest sentral for den blå gruppen. Endelig er node 11 en bronode som er avgjørende for flyten av informasjon mellom de to grupper. 25  
30

Anta videre at sikkerheten for dette lille systemet er av interesse. Da kan man umiddelbart identifisere node 9, 13 og 11 som de som mest øyeblikkelig trenger beskyttelse fra det som truer systeminngangene. Noder 9 og 13 skal beskyttes fordi de er sentre for sine regioner: Hvis de infiseres er det stor sannsynlighet for deres hele regioner også vil bli infisert. 35

Videre kan man gi node 9 en høyere prioritet for beskyttelse enn node 13, siden dens region er større.

Endelig fortjener node 11 ekstra beskyttelse, siden hvis  
5 den kan gjøres immun mot truslene, vil disse trusler ikke ha noen ferdig kanal for spredning fra en region til en annen.

Legg merke til at bruk av Regel 2 ikke peker ut  
10 noen grensenoder for spesiell beskyttelse - selv om node 11 klart spiller en viktig rolle i å koble de to regioner. Imidlertid vil Regel 2 identifisere *linken* mellom 11 og 13 som en brolink. Den åpenbare konsekvens av dette er at nodene på hver ende av hver brolink fortjener spesielle  
15 beskyttende tiltak.

Problemet kan snus på hodet, ved å gi administratoren problemet med å *spre ønsket informasjon* over dette samme lille nettverk. Analysen foreslår da en effektivstrategi  
20 for å gjøre dette: Man stater med sentrene (node 9 og 13), og arrangerer for at den ønskede informasjon skal kringkastes derifra.

Det er selvfølgelig å forvente at avstandsregelen og  
25 bratteste-oppstignings-regelen vil gi motsatte resultater for noen noder. Et viktig punkt å hente fra Figurene 2 til og med 7 er at det generelle kvalitative bilde er nokså ufølsomt for valget av regel for å definere regioner. Man kan forvente at dette er tilfelle for de fleste grafer.  
30 Valget av sentre er uavhengig av hvilken regel som brukes; og disse sentre eksisterer i sin tur nøyaktig fordi de ligger pi en region av grafen som har noe "vekt" - det vil si et antall noder som er bedre koblet til hverandre enn til deres "omgivelser". Kort sagt vil de distinkte regler,  
35 som tilsynelatende definerer regioner, faktisk avvike hovedsakelig ifølge hvor de plasserer grensene mellom regioner - mens regionene selv er heller stabile objekter.



**Sammenfatning av definisjonene av roller og regioner i nettverk.**

- 5 Grunnkriteriet for å definere en region (og dens senter) har vært veltilknyttethet, som målt av den "glatte" grafefunksjon, egenvektorsentralitet eller EVC. I tillegg til å definere naturlige klynger av en grafe, tildeler vår fremgangsmåte en unik rolle til hver node i grafen.
- 10 De to regler som definerer regioner gir kvalitativt lignende bilder for grafestrukturen som et hele, med nokså forskjellige bilder når det gjelder hvilke roller for nodene som er til stede i analysen.
- 15 Det vil si at regel 1 - som assosierer noder med regioner basert bare på deres avstand, i korteste banehopp, fra sentre - plasserer et betydelig antall noder i grensesettet. Disse noder kan i sin tur plasseres i to distinkte roller: Bronoder og dinglere (se Figur 2). Regel
- 20 2 holder seg nærmere til den "topologiske" ånd av fremgangsmåten som beskrevet innen foreliggende søknad, og assosierer noder med sentre til hvilke de er linket med bratteste bane oppstigning. Denne regel vil normalt (i fravær av spesiell symmetri) plassere ingen noder i
- 25 grensesettet - slik at, med regel 2, blir de to roller i grensesettet (bronoder og dinglere) vesentlig ekskludert, og alle nodene er enten sentre i en region, eller medlemmer av en region.
- 30 Man kan forestille seg andre regler for å definere regioner. Hovedaspektet av fremgangsmåten ifølge foreliggende oppfinnelse er å identifisere sentre først, og deretter la regionene "vokse" utover fra disse sentre. Begge reglene i samsvar med foreliggende oppfinnelse passer
- 35 dette bilde. Girvan/Newman-fremgangsmåten tillater en hierarkastisk dekomponering av en grafe, ved å bryte klynger opp i sub-klynger osv. En tilsvarende hierarkastisk dekomponering kan også utføres ifølge foreliggende

oppfinnelse, ved å eliminere grensenoder og linker, og anvende analysen ifølge foreliggende oppfinnelse på de resulterende isolerte regioner. Videre kan regler defineres basert på de foreliggende analysemetoder. I et meget enkelt eksempel kan man tildele rollen "Kanten av regionen" til de noder som er koblet til grenseelementer (noder eller linker). En annen type Kant-rolle kan tildeles til de noder som er "fjernest" fra senteret, men ikke linket til noe grenseelementer.

10

### **Bruksområder**

I det etterfølgende er det gitt bruksområder for fremgangsmåten og systemet ifølge foreliggende oppfinnelse. Klart kan både høyt sentrale noder, og broer (linker eller noder) peker ut som å fortjene ekstra oppmerksomhet og pleie for å hindre spredning av ødeleggelse. De høyt sentrale noder vil mest sannsynlig infisere deres regioner; og broene må i sin tur voktes slik at infeksjonen ikke spres fra en region til andre. Derfor vil det være praktisk å immunisere visse elementer, og å sikre at enhver infeksjon blir isolert til en enkelt region. For større regioner vil det også være praktisk å immunisere de mest sentrale noder i hver region - og selvfølgelig prioritere regionene med det største antall noder. Noen eksempler slik som meget veltilknyttede "peer-to-peer" systemer er på den annen side vanskelig å beskytte, fordi de er for godt tilknyttet. Dette betyr at det er mange noder i hver region med grovt sett den samme sentralitet, og at det er mange broer mellom regioner (for de tilfelle hvor det er mer enn en region).

Systemet og fremgangsmåten kan anvendes på mange andre typer grafer - i prinsippet på enhver grafe som er ikke-rettet. Fremgangsmåten kan enkelt modifiseres - som beskrevet i den første utførelse - til også å tillate vekter (annet enn 0 og 1) for linkene mellom noder. Fremgangsmåten og systemet ifølge foreliggende oppfinnelse

vil vise seg å være nyttig i analysen av sosiale nettverk - som (igjen) kan ha en (positiv) styrke assosiert med hver link.

5 Når en innovasjon - et nytt produkt eller tjeneste - introduseres i en befolkning, følger diffusjonen av innovasjonen et typisk mønster. Innovasjonen oppdages vanligvis av en liten gruppe tidlige brukere, og etter en stund, avhengig av de tidlige brukernes godkjennelse, vil  
10 opinionslederne (eller ledende brukere) adoptere innovasjonen. dette er det kritiske punkt i diffusjonsprosessen, fordi adopteringen av innovasjonen av majoritetsbefolkningen vanligvis avhenger av aksepteringen av opinionslederne. [6]. Med andre ord tar adopsjonen av en  
15 innovasjon av når opinionslederne eller sosiale nettverksnav godkjenner og adopterer innovasjonen.

Fremgangsmåten som beskrevet ved utførelsen og dens medfølgende eksempler av foreliggende oppfinnelse, bruker  
20 en modifisert naboskapsmatrise, basert på flytdata, til å beregne et sentralitetsmål for hver node i et sosialt nettverk. Denne sentralitetsindeks tillater de mest sentrale noder i det sosiale nettverk som denne naboskapsmatrise representerer å plukkes ut. Disse noder -  
25 nettverksnavne - er i sosiale nettverkstermer, opinionsledere. De er derfor gode mål for spredning av informasjon osv., fordi de kan potensielt bidra til akselerert diffusjon av slik informasjon. En åpenbar anvendelse av fremgangsmåten er derfor innen området  
30 innovasjonsdiffusjon.

I innledningsdelen har det blitt vist til epidomlogi, telekommunikasjon, datakom, elektriske strømsystemer osv. Det kan legges til at resultatet av analysen ifølge  
35 foreliggende oppfinnelse videre har et vidt område av applikasjoner. Et eksempel er planlegging av tidstabeller innen transport, eller sendings- og distribusjonssystemer. Ved å analysere trafikkflyt in et nettverk av veier eller

et jernbanesystem, kan den beste timing for distribusjon finnes for å unngå trafikkmetning. Tilsvarende er planlegging av trafikkrouting innen telekom eller datakom en åpenbar applikasjon av foreliggende oppfinnelse, fordi  
5 fremgangsmåten enkelt kan identifisere metningspunkter eller gode ruter. Videre på et mer mikroskopisk nivå kan den brukes innen design av datamaskiner, for å analysere intern trafikk og derved optimalisere dens komponenter og dens busser. Det siste er særlig nyttig innen området  
10 parallellprosessering, for å redusere trafikken mellom prosessorer/datamaskiner.

Legg merke til at mens det i foregående har det blitt tilveiebragt en detaljert beskrivelse av foreliggende  
15 oppfinnelse, er det å forstå at ekvivalenter skal inkluderes innen omfanget av oppfinnelsen som definert i kravene. Den detaljerte beskrivelse har i stor utstrekning behandlet teorien bak foreliggende oppfinnelse, imidlertid har bruk av disse teorier et bredt utvalg av bruksområder,  
20 forutsatt at grafene er ikke-rettede.

På en generell basis er derfor fremgangsmåten ifølge foreliggende anvendbar innen et vidt utvalg av områder, og den kan anvendes for å løse problemer innen disse områder.  
25 Andre fordelaktige utførelser av foreliggende oppfinnelse vil fremgå av de vedføyde uselvstendige krav.

#### **Forkortelser og referanser**

30 1. G. D. BATISTA, P. EADES, R. TAMASSIA, AND I. G. TOLLIS, *Graph Drawing: Algorithms for the Visualization of Graphs*, Prentice Hall, Upper Saddle River, New Jersey, 1999.

2. P. BONACICH, *Factoring and weighting approaches to status scores and clique identification*, *Journal of*  
35 *Mathematical Sociology*, 2 (1972), pp. 113-120.

3. M. BURGESS, G. CANRIGHT, AND K. ENGØ, *A graph theoretical model of computer security: from file access to social engineering*, International Journal of Information Security, (2003). submitted for publication.
- 5
4. G. CANRIGHT, K. ENGØ-MONSEN, AND Å. WELTZIEN, *Multiplex structure of the communications network in a small working group*, Social Networks - An International Journal of Structural Analysis, (2003). submitted for publication.
- 10
5. M. GIRVAN AND M. NEWMAN, *Community structure in social and biological networks*, Proc. Natl. Acad. Sci. USA, 99 (2002), pp. 8271-8276.
- 15
6. E. M. ROGERS, *Diffusion of Innovations*. Free Press, Fifth Edition, 2003.
7. J. M. KLEINBERG, *Authoritative sources in a hyperlinked environment*, Journal of the ACM, 46 (1999), pp. 604-632.
- 20
8. M. NEWMAN, *The structure and function of complex networks*, SIAM Review, 45 (2003), pp. 167-256.
9. A. Y. NG, A. X. ZHENG, AND M. I. JORDAN, *Stable algorithms for link analysis*, in Proc. 24th Annual Intl. ACM SIGIR Conference, ACM, 2001.
- 25
10. A. ORAM, ed., *Peer-to-peer: Harnessing the Power of Disruptive Technologies*, O'Reilly, Sebastopol, California, 2001.
- 30
11. L. PAGE, S. BRIN, R. MOTWANI, AND T. WINOGRAD, *The pagerank citation ranking: Bringing order to the web*, tech. report, Stanford Digital Library Technologies Project, 1998.
- 35
12. R. PASTOR-SATORRAS AND A. VESPIGNANI, *Epidemic spreading in scale-free networks*, Phys. Rev. Lett., 86 (2001), pp. 3200-3203.

13. T. H. STANG, F. POURBAYAT, M. BURGESS, G. CANRIGHT, K. ENGØ, AND Å. WELTZIEN, *Archipelago: A network security analysis tool*, in Proceedings of The 17<sup>th</sup> Annual Large  
5 Installation Systems Administration Conference (LISA 2003), San Diego, California, USA, October 2003.

14. G. H. GOLUB AND C. H. VAN LOAN, *Matrix Computations*. The Johns Hopkins University Press, Second Edition, 1989.

**P a t e n t k r a v**

1. Fremgangsmåte for å bestemme et nettverks evne til å spre informasjon eller fysisk trafikk, idet nevnte nettverk  
5 omfatter et antall nettverksnoder sammenknyttet med linker, idet fremgangsmåten omfatter trinnene:

- kartlegging av topologien av et nettverk,
- beregning av en verdi for linkstyrke mellom nodene,
- beregning av en Egenvektor Sentralitetsindeks for alle  
10 noder basert på nevnte linkstyrkeverdier,
- identifisering av noder som er lokale maksima av Egenvektor Sentralitetsindeksen som senternoder,
- gruppering av nodene i regioner som omgir hver identifisert senternode,
- 15 • tildeling av en rolle til hver node fra dens posisjon i en region, som senternoder, regionmedlemsnoder, grensenoder, bronoder, dinglenoder,
- måling av nettverkets mottagelighet for spredning, basert på antall regioner, deres størrelse, og hvordan  
20 de er sammenknyttet

k a r a k t e r i s e r t v e d

- tildeling av rollen som regionmedlemsnoder i en gitt region til alle noder for hvilke den bratteste oppstigningslinkbane i topologikartet terminerer unikt  
25 i senternoden for denne region.

2. Fremgangsmåte ifølge krav 1,

k a r a k t e r i s e r t v e d

30 beregning av nevnte linkstyrke-verdi ved å telle antall forskjellige typer bindinger ethvert par av noder benytter for deres samvirkning, og bruke antall bindinger som et mål for linkstyrke.

35 3. Fremgangsmåte ifølge krav 1,

k a r a k t e r i s e r t v e d

beregning av nevnte link styrke verdi ved å måle trafikken mellom enhver to noder, og bruke målet på trafikken som et mål på linkstyrke.

5 4. Fremgangsmåte ifølge krav 1,  
karakterisert ved  
beregning av nevnte linkstyrke-verdi ved å måle trafikken mellom hvert par av noder for hver forskjellig type binding, dele trafikkmengden i hver type binding med den  
10 totale trafikk for denne type binding, benytte summen av de resulterende fraksjoner som et mål på bindingsstyrke.

5. Fremgangsmåte ifølge et av kravene 1 - 4,  
karakterisert ved  
15 organisering av nevnte linkstyrkeverdier i en matrise, naboskapsmatrisen, og beregne Egenvektor Sentralitetsindeksen som hoved-egenvektoren i nevnte naboskapsmatrise.

20 6. Fremgangsmåte ifølge krav 1,  
karakterisert ved  
tildeling av rollen som grensenoder til alle noder som ikke har noen unik assosiasjon til noen senternode.

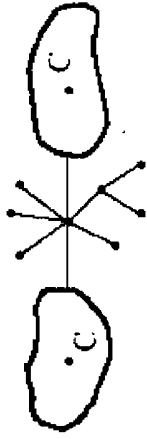
25 7. Fremgangsmåte ifølge krav 1,  
karakterisert ved  
tildeling av rollen som bronoder til alle grensenoder som ligger på i det minste en ikke-selv-gjenfindende linkbane  
30 som forbinder to senternoder.

8. Fremgangsmåte ifølge krav 1,  
karakterisert ved  
tildeling av rollen som dinglenoder til alle grensenoder,  
35 og ligger på i det minste en ikke-selv-gjenfindende linkbane som forbinder to senternoder.

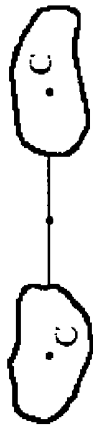


9. Anvendelse av fremgangsmåten ifølge ethvert av kravene 1 - 8 for å hindre spredning av virus eller skadelig informasjon i et nettverk.
- 5 10. Anvendelse av fremgangsmåten ifølge ethvert av kravene 1 - 8 for å forbedre spredningen av informasjon i et nettverk.
- 10 11. Anvendelse av fremgangsmåten ifølge ethvert av kravene 1 - 8 for planlegging av arkitekturen i et nettverk, for å forbedre robusthet og/eller sikkerhet og/eller kommunikasjonseffektivitet i nevnte nettverk.
- 15 12. Anvendelse av fremgangsmåten ifølge ethvert av kravene 1 - 8 for planlegging av arkitekturen i et kraftnett for å forbedre robustheten av nevnte nettverk.
- 20 13. Anvendelse av fremgangsmåten ifølge ethvert av kravene 1 - 8 for planlegging av et distribusjonsnett for gods.
14. Anvendelse av fremgangsmåten ifølge ethvert av kravene 1 - 8 for planlegging av et transportnett.

1/7

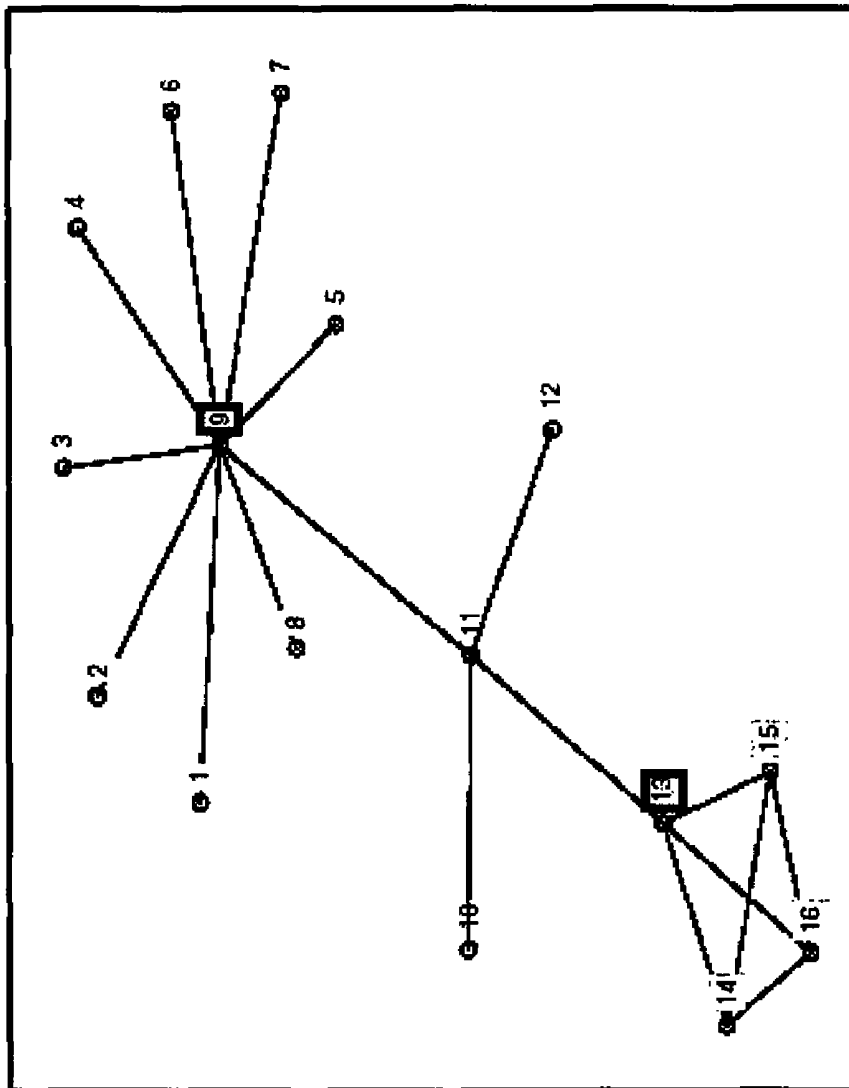


(høyre)

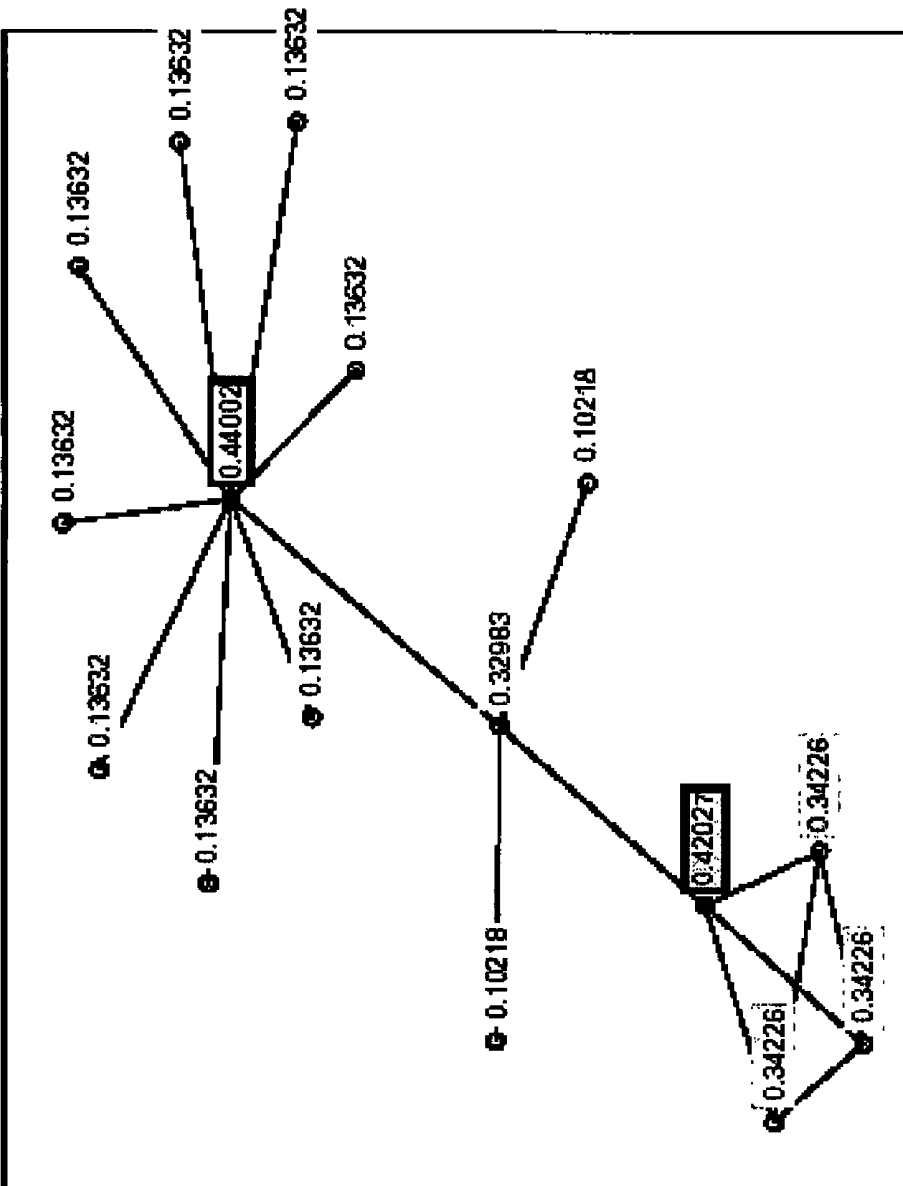


(venstre)

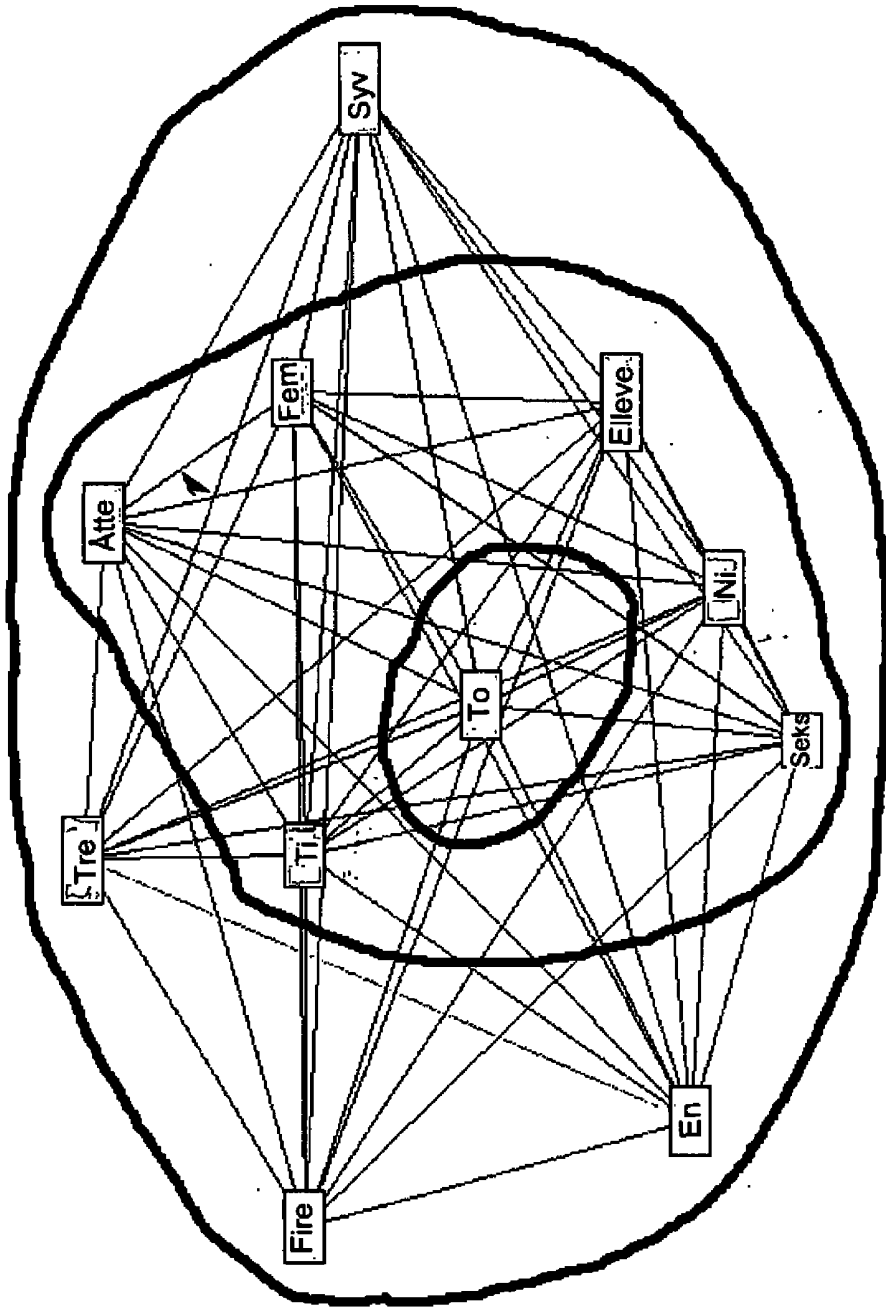
Figur 1



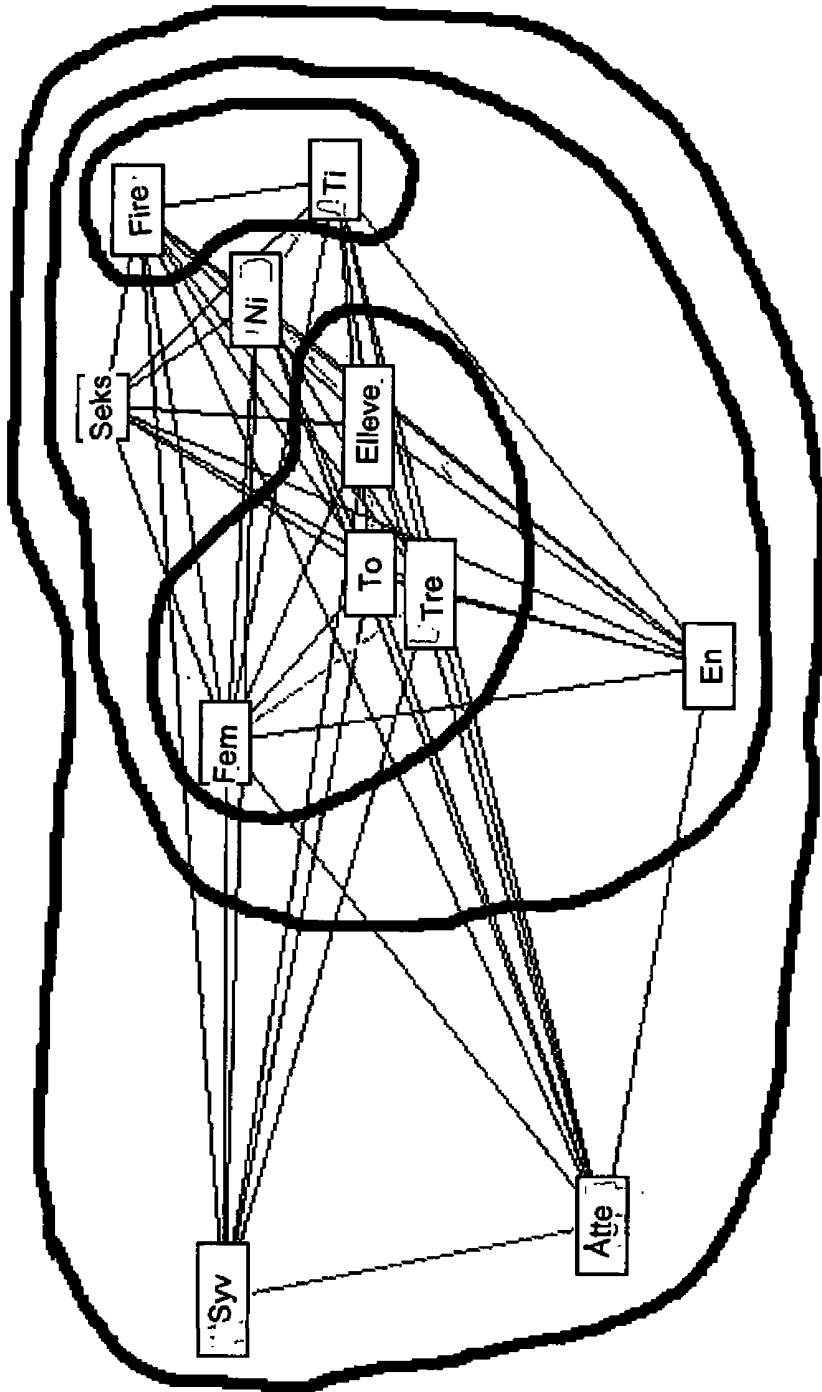
Figur 2



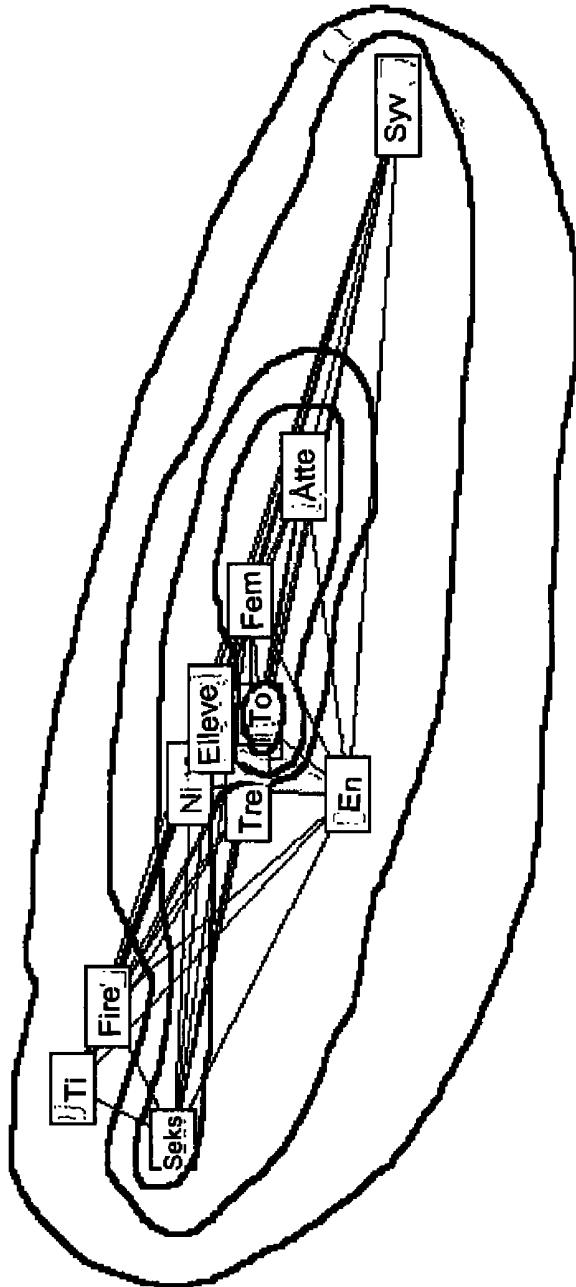
Figur 3



Figur 4



Figur 5



Figur 6

