

XIII. Erdélyi Tudományos Diákköri Konferencia

Korszerű adatelemzési módszerek az alvás-elektrofiziológiában

Készítette:

Kacsó Ágota

Papp István

Suciaghi Róbert

Babeş-Bolyai Tudományegyetem

Fizika Kar, III. év

Témavezető:

Lázár Zsolt-József

egyetemi adjunktus

Babeş-Bolyai Tudományegyetem

Fizika Kar

Kolozsvár

2010

Tartalomjegyzék

| | |
|---|-----------|
| 1. Elektroencefalográfia, alvástudatás | 5 |
| 1.1. Elektroencefalográfia | 5 |
| 1.2. Alvástudatás | 6 |
| 1.3. Mennyiségi adatelemzési módszerek | 7 |
| 1.4. Adataink | 8 |
| 1.5. Számítógépes módszereink | 8 |
| 2. Nemlineáris módszerek | 9 |
| 2.1. Beágyazás | 10 |
| 2.1.1. Hamis szomszédság | 11 |
| 2.1.2. Az időkésés meghatározása | 12 |
| 2.2. Fluktuáció elemzés | 12 |
| 2.3. Komplexitás | 13 |
| 2.4. Determinizmus és előrejelezhetőség | 16 |
| 2.4.1. Előrejelezhetőség | 16 |
| 2.4.2. Egyszerű nemlineáris előrejelző algoritmus | 16 |
| 2.4.3. Egyszerű nemlineáris zaj szűrés | 17 |
| 3. Független komponens elemzés | 19 |
| 3.1. Mi a független komponens analízis? | 19 |
| 3.2. Az ICA tulajdonságai | 21 |
| 3.3. Az ICA szerepe az alvás-elektrofiziológiában | 21 |
| 4. Hálózatelemzési módszerek | 25 |

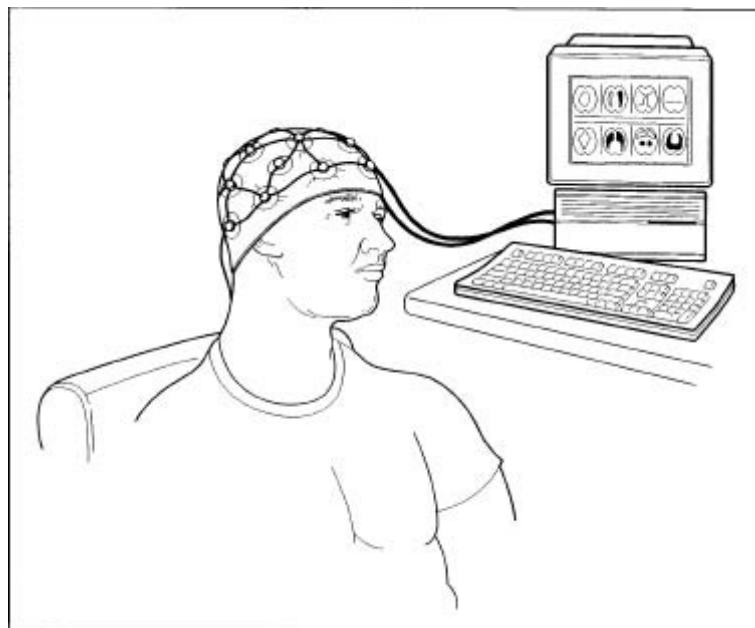
| | |
|--|-----------|
| 4.1. A modern hálózatelmélet alapjai | 25 |
| 4.2. Idegtudományi alkalmazások | 27 |
| 4.3. Funkcionális hálózatok az alváskutatásban | 28 |
| 5. Optimalizált számítógépes infrastruktúra | 33 |
| 6. Következtetés és kitekintés | 35 |
| Irodalomjegyzék | 37 |

1. fejezet

Elektroencefalográfia, alváskutatás

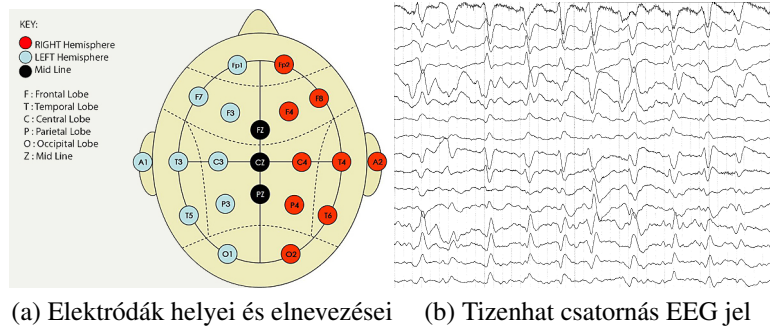
1.1. Elektroencefalográfia

Az elektroencefalográfia (EEG) az agyi tevékenység egy nem invazív mérési módja, a legelterjedtebb agyi képalkotói eljárás. A jó vezetőképességű ragasztóval a fejbőrre rögzített elektródák által felfogott elektromos potenciál (1.1 ábra), milliárd és milliárd idegsejt tevékenységének átlagos értéke. Szemben a funkcionális mágneses magrezonancia képalkotási módszerekkel (fMRI) melyek sokkal nagyobb térbeli felbontást biztosítanak, az EEG előnye a magas időbeli felbontás és a nagyságrendekkel alacsonyabb költség. Alapja, hogy az idegsejtek elektromos impulzusok segítségével kommunikálnak



1.1. ábra. Elektroencefalográfiai mérés

egymással. Egy adott agyterületről érkező impulzusokat összegyűjtve és felerősítve vizsgálni lehet a keletkező agyhullámokat és következtetni lehet az azokat generáló agyi területek állapotára. A 2 - 130 elektródáról begyűjtött EEG jel (1.2a ábra), több komponensből álló, periodikus változásokat mutató görbesereg (1.2b ábra). Az maximum $100\mu V$ (mikrovolt) amplitúdójú analóg jelet 128 - 1024 hertzen mintavételezik és digitalizálják.



Az jel különböző forrásból származó és eltérő szabályosságokat követő összetevők keveréke. Ezen komponenseket amplitúdójuk és frekvenciájuk alapján osztályozzák. Az EEG megváltozása diagnosztikai jelentőségű. Fontos klinikai eszköz bizonyos betegségek követésére, kimutatására és kezelésére. Agytumorok, rohamok, epilepszia, fertőző betegségek, elbutulás, különböző feji sérülések, drog túladagolás, alvás zavarok, és agyhalál, csak néhány az orvosi problémák közül ahol megjelenhet az EEG. Emellett állatokon is sikeresen alkalmazzák.

Az újabban kifejlesztett EEG regisztrálási és értékelési módszerek segítségével lehetőség nyílik az „agy feltérképezésére”, vagyis az agy működés közbeni állapotainak leírására.

1.2. Alváskutatás

Az EEG méréseket két nagy osztályba sorolhatjuk. A *esemény kiváltott potenciál esetében* ébrenléti állapotban különböző külső ingerekre adott válaszból vonnak le következtetéseket az agyi funkciókra vonatkozóan. Az *alvás EEG* esetén az adatokat alvás közben gyűjtik be a kísérleti alanyról. Az alvás EEG előnye, hogy az agy alvásban a legelszigeteltebb külső tényezőktől, úgy mint a beszűrődő zaj, fényvillanások stb. Ilyenkor az agy „saját hangját hallani”, ami egyedüli módja például különböző patológiák diagnosztizálásának illetve megértésének. Ébrenlétkor sokkal érzékenyebbek vagyunk a környezeti változásokra és ezek igen nagy intenzitással megjelennek az EEG-ben is, elnyomva a személyre vonatkozó specifikusabb összetevőket.

Teljes éjszakai mérés esetén 256 Hz mintavételi frekvencián csatornánként $256 \times 9 \times 3600 \approx 10^7$ érték kerül rögzítésre. Egy sokcsatornás mérés esetén 32-64 bites számábrázolással történő rögzítést követően az elemzésre váró állomány mérete gigabyte-os nagyságrendű.

Az adat háromszintű elemzésen esik keresztül. Első lépésben egy alvásszakértő vizuális úton megjelöli és/vagy kivágja az ún. *műtermékes* szakaszokat, melyeket szennyezi valamilyen külső zavaró tényezőkre adott válasza az agynak, úgy mint hanginger, például leejtett kulcs, vagy valamely motoros tevékenység, például a kísérleti személy önkéntelen testhelyzetváltoztatása alvás közben. Ilyenkor általában a jelből kiszűrik az 50Hz-s frekvenciát, amivel a szobában működő elektromos készülékek szennyezik az EEG jelet. Az alváskutató beazonosítja az alvás különböző stádiumainak – ébrenlét, REM (Rapid Eye Movement), NREM (Non-REM) 1, 2, 3 – kezdeti és végpontjait. Ezek összességét *hipnogrammnak* nevezik. A műtermékmentesítést és hipnogramkészítést követően az adatot *mennyiségi elemzésnek* vetik alá. Ennek során egy vagy több változót számolnak ki a jelből. Ezt a folyamatot is megismétlik minden egyes kísérleti személyre. Az így kapott adathalmazból *statisztikai elemzéssel* szűrik ki a végső következtetéseket.

1.3. Mennyiségi adatelemzési módszerek

A XX. század második feléig kizárólagos módon lineáris elemzési módszereket használtak. Ez jobbra a jel spektrumának és a csatornák közötti koherenciának a mélyreható tanulmányozásában merült ki. Az utóbbi évtizedekben, a matematika, számítástechnika és interdiszciplináris kutatás fejlődésével újabb módszerek sokasága jelent meg, és az EEG újra előtérbe került az agyi képalkotásban.

Egy lehetséges stratégia, hogy az agyat úgy tekintjük, mint egy komplex dinamikai rendszer, és olyan más tudományterületeken keresünk új megközelítési módokat, amelyek ugyancsak komplex rendszerek tanulmányozásával foglalkoznak. A matematika és fizika három kutatási ága különösen alkalmasnak bizonyult az általános komplex hálózatok tanulmányozására: (i) nemlineáris dinamika és a kapcsolódó területei, mint például a szinergetika (ii) statisztikus fizika amely a fázisátalakulások egyetemes jelenségével, valamint a skálázási móddal foglalkozik, és (iii) a hálózatok modern elmélete, amely a gráfelméletből alakult ki.

Mivel az EEG jelet több periódikusan változó élettani folyamat befolyásolja, úgy mint a szív működés vagy izzadás, ezek kiszűrése egy külön feladatkörre fejlődött.

1.4. Adataink

A fejlesztett módszereket két Budapesten folyó kísérletből származó adatokra alkalmaztuk az illető csoportok kérésére. Az egyik kísérlet enyhén autista (aspergeres) gyerekeken végzett alvástanulmány része. Ez 31 személy, 17 kísérleti (aspergeres) és 14 kontrol, teljes éjszakás alvásának, egyenként tízcsatornás, 256Hz-es, műtermékmentesített jele. Ezekre „vadaskerti adatok” néven fogunk hivatkozni a továbbiakban. A második kísérlet adatai, a továbbiakban „MICS”, veseátültetésen és a megfelelő kezeléssel átesett személyek vizsgálatából származik. Ez három EEG csatorna, további két fülmögötti referencia, két szemcsatorna (elektrookulogram (EOG)) és egy elektrokardiogram (EKG) csatornából áll.

1.5. Számítógépes módszereink

A jelen dolgozatban elért eredmények eléréséhez kizárólagos módon a Python programozási nyelvet használtuk. A népszerű és teljes értékű programozási nyelvhez matematikai és adatfeldolgozás könyvtárak sokasága tartozik. Ez a Matlabszerű magas szintű programozási nyelv más gondolkodásmódot követelnek. A kihívás az volt, hogy minimálisan vegyük igénybe a számítógép erőforrásait, kis programokat írjuk amelyek nem tartalmazznak az időtengelyt hosszanti irányba bejáró ciklusokat, kivételt képez az az eset, amikor a jelet ablakokra, ún. *epoch*-okra osztjuk, mert ez többszörösen lerövidíti a bejárás utat.

A futtatás jórészéhez a BBTE Fizika Karának számítógépes klaszterét használtuk a célra fejlesztett szkripteken keresztül. Erre bővebben a 5 fejezetben térünk ki.

2. fejezet

Nemlineáris módszerek

Nagyon sok érdekes esemény a természetben nemlineáris esemény, ezért is a lehetősége egy determinisztikus káoszelméletnek befolyásolta a tudósok gondolkodását a tudomány több ágazatában is, a legerősebb kapcsolat a természet és a káosz elmélet között az idősor elemzésében rejlik. Több jelenség leírható mint egy dinamikai rendszer például a banki kamat, a Föld lakosságának növekedése, időjárás, a Nap és más bolygók kémiai folyamatai, a tőzsde index változása, vagy például az áramkörök. Annak ellenére, hogy ezek teljesen különböző rendszerek mégis leírhatók mint dinamikai rendszerek. A dinamikai rendszerek változatosak, megkülönböztethetünk lineáris és nem lineáris rendszereket. A rendszer lineáris, ha minden függvény ami leírja a rendszer viselkedését lineáris. Ezért lineáris kapcsolat van az okok és hatások között. Egy nemlineáris rendszernél ez nem törvényszerű vagyis kis ok akár nagy hatást is kiválthat.

A nemlineáris dinamikát 1985 óta alkalmazzák az agykutatásban és egy nagyon aktív kutatási területté vált. Az idegtudományokban való alkalmazásának egyik eredménye egy új modell az epilepsziás rohamok leírására, amely észleli, és esetleg előre is jelzi a roham bekövetkezését. Újabban azonban inkább az agy különböző területei közötti nemlineáris kölcsönhatások tanulmányozására koncentrálnak.

Egy módja a nemlineáris idősor analízisnek az EEG jelek időfüggvényének megfigyelése, dimenziójának meghatározása, Lyapunov exponens és entrópia kiszámolása. A kiindulási pont a dinamikai rendszerekhez hasonló, azt feltételezzük, hogy egy következő állapot a kezdeti állapot valamilyen függvénye. Hogyha ismernénk a rendszert meghatározó összes függvényt a fentiekben felsorolt összes eredményt megkaphatnánk, de sajnos a kórházi mérések eredménye nem egy egyenletrendszer hanem

egy EEG idősor. A nemlineáris módszerek lehetőséget adnak, hogy ebből az idősből adatokat nyerjünk az eredeti sokdimenziós rendszer tulajdonságaira vonatkozóan.

2.1. Beágyazás

Az első és legfontosabb lépés a nemlineáris analízisnél a jel átrajzolása egy vagy néhány idősből a megfelelő dimenzióba. Két különböző beágyazási módszer használatos: i.) időkésés beágyazás, ii.) tér beágyazás [4].

Az első esetben egy egyszerű idősből indulunk ki és ezt fogjuk átépíteni m dimenzióba (lásd a 2.1 ábrát). Legyen ez az s_n idősor valamilyen mennyiség skaláris méréseinek halmaza, azaz a rendszer jelenlegi állapotának függvénye, amelyet azonos időközönként olvasunk le. Kiveszünk egymástól τ távolságra elhelyezkedő m elemet az idősorunkból és ezeket egy m dimenziós vektor megfelelő koordinátáinak fogjuk fel. Ezt megisméteve, a jel végéig eredményként egy m dimenziós jelet kapunk:

$$\mathbf{s}_n = (s_{n-(m-1)\tau}, s_{n-(m-2)\tau}, \dots, s_{n-\tau}, s_n)$$

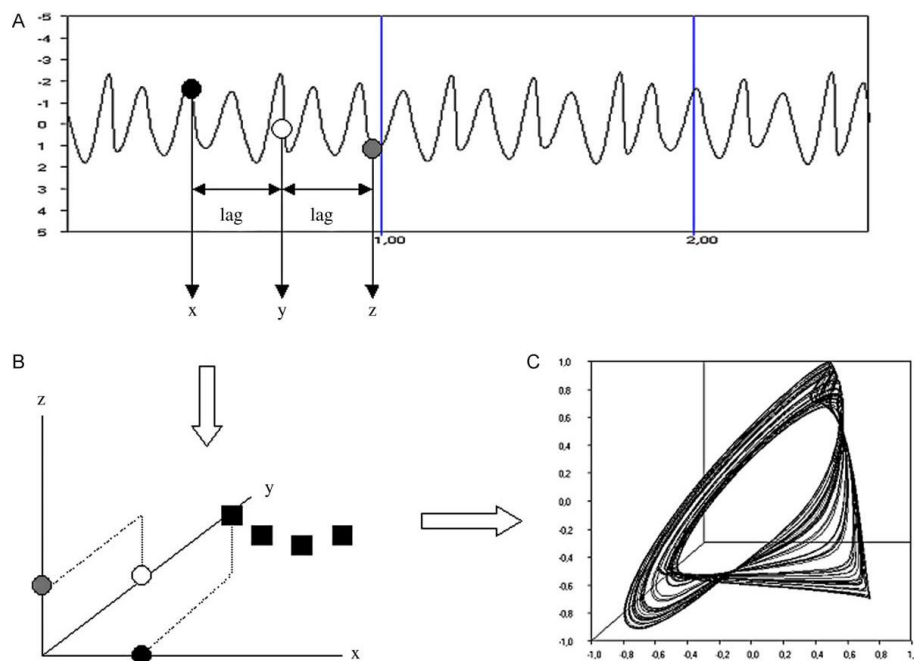
A legtöbb esetben mindkét paraméter – m és τ – pontos meghatározása elengedhetetlen. Túl nagy m érték használata redundanciát eredményez és természetesen lelassítja az algoritmus működési sebességét. Túl kis τ -t használva a jel túlkorrelál, túl nagy τ -t használva teljesen független vektoraink lesznek. Ezek a problémák különösen zajos jel esetén igényelnek több figyelmet.

Térbeágyazás esetén a vektor m ordinátáit az m idősor eredményének vesszük, ezt ismételve egy sor vektort kapunk eredményként. Ebben az esetben a beágyazási dimenzió m egyenlő a megfelelő csatornákkal, amelyeket használunk, hogy újraépítsük a jelünk.

A τ időkésés kiszámítására használható az önkorreláció illetve az ún. kölcsönös információ. Mindkét mennyiség a jelnek önmagával való eltolásának segítségével számolandó ki és hagyomány szerint az így kapott $c(\tau)$ illetve $I(\tau)$ függvények első minimumhelye. Sikerült kiszámolni, hogy a vadaskerti EEG adatra és az értéke megközelítőleg 1.76 másodperc.

A dimenzió meghatározása már nehezebb feladat. Ún. *hamis szomszéd* módszerrel a háttérdinamika folytonosságából kiindulva keressük azt a dimenziót, amiben minimális az olyan pontok száma, melyek az m dimenziós térben szomszédosok, de az (egydimenziós) jelben nem azok.

Az általunk készített modul pillanatnyilag csak többszörös próbálgatással keresi az optimális di-



2.1. ábra. Időbeágyazás. A. kiválasztjuk a ordináták értékeit a megfelelő időkéssel B. a megfelelő pontokat elhelyezzük az új többdimenziós térben C. a jel viselkedése az új dimenzióban Forrás: [3]

menziót.

2.2. Fluktuáció elemzés

A fraktálság egy gyakran előforduló tulajdonsága a természetnek, de megfigyelhető idősorokban is, ezért a fluktuáció elemzés (DFA)¹ egy hasznos eszköz lehet az agy működésének a leírásában, de használható sztochasztikus analízisben, káosz elméletben. Ez a módszer a függvény önhasonlóságát vizsgálja. Az eljárás eredményeképpen kapott érték rokon az ún. Hurst-exponenssel, azzal a különbséggel, hogy a DFA alkalmazható olyan függvényekre, amelyek dinamikája nem helyhez kötött (időben változó). Az alváskutatásban a Hurst-exponens jól jellemzi a különböző alvásstádiumokat, ezért potenciálisan automatikus hipnogramkészítésre is alkalmazható [5].

Az $x_t, t = \overline{1, N}$ idősornak az $m(x)$ középértéktől való eltérését felösszegezve létrehozuk az

$$X_t = \sum_{i=1}^t (x_i - \bar{x}_i)$$

újabb idősort. Ezek után $X_t - t L$ hosszúságú időablakokra osztjuk és lineáris regressziót végzünk

¹ Detrended Fluctuation Analysis

minimalizálva az illesztő egyenestől számított távolságok négyzetösszegét. A kapott értéket különböző L ablakméretekre kiszámolva egy $F(L)$ függvényt kapunk, mely $F(L) \propto L^\alpha$ típusú viselkedése esetén az idősor önhasonlóságról beszélhetünk. Az α exponens hasznos információkat árul el a jelről:

$\alpha < 1/2$ anti korreláló

$\alpha \sim 1/2$ nem korrelált, fehér zaj

$\alpha > 1/2$ korrelál

$\alpha \sim 1$ $1/f$ -zaj, rózsaszín zaj

$\alpha > 1$ nem stacionárius

$\alpha \sim 3/2$ Brown zaj

Az EEG esetén az exponens a hosszú távú korrelációt határozza meg,

ha az idősor tulajdonságai invariánsak az újraszámítás után az mondhatjuk, hogy az idősor önhasonló. Az α exponens növekszik 1,2 és 3,4 alvási stádiumokban, az ébrenléti állapothoz képest, de csökken a gyors szemmozgás (REM) állapotban. Fontos megjegyezni, hogy az α exponense az EEG-nek növekszik amint az agy mély alvásba kezd, ami valószínűleg Brown zajnak köszönhető, ebből arra következtetünk, hogy az agy kevésbé aktív mély alvás közben [6].

A leprogramozott algoritmus helyességét Brown zajra teszteltük, ezek után alkalmaztuk a vadaskereti EEG jelre, feltételezve, hogy különbségek észlelhetők a két kísérleti csoport között. Ez utóbbinak a megerősítéséhez még szükséges a statisztikai elemzés.

2.3. Komplexitás

A térbeágyazás egy jól használható elektrofiziológiai alkalmazása a jel komplexitásának a mérése [7]. Legyen a bemenet K csatornáról beolvasott szimultán mérések eredménye, melyeket az $\mathbf{u}_n = (u_n^1, u_n^2, \dots, u_n^K)$, $n = 1, 2, \dots$ feszültségvektorokban tárolunk. Ezek egy K dimenziós térben lejátszódó dinamikát határoznak meg.

Ennek tanulmányozására osszuk a jelet T hosszúságú, egyenként N vektorból álló ún. epoch-ra, és minden epochban hozzuk nulla átlagra, azaz $\mathbf{u}_n \rightarrow \mathbf{u}_n - \langle \mathbf{u} \rangle_{t, t+T}$. A továbbiakban minden esetben egy adott epochra vonatkoztatunk.

Az \mathbf{u} vektor változása során egy pályát ír le a K dimenziós térben. Kiszámolva az

$$m_0 = \frac{1}{N} \sum_n \|\mathbf{u}_n\|^2, \quad \text{és} \quad (2.1)$$

$$m_1 = \frac{1}{N} \sum_n \left\| \frac{\Delta \mathbf{u}_n}{\Delta t} \right\|^2, \quad \text{ahol} \quad \Delta \mathbf{u}_n = \mathbf{u}_n - \mathbf{u}_{n-1}. \quad (2.2)$$

mennyiségeket a K dimenziós „mozgás” három globális jellemzőjét vizsgáljuk:

$$\Sigma = \sqrt{\frac{m_0}{K}}, \quad \text{átlagos térerősség mértéke} \quad (2.3)$$

$$\Phi = \frac{1}{2\pi} \sqrt{\frac{m_1}{m_0}}, \quad \text{változás mértéke}. \quad (2.4)$$

és végül az Ω komplexitást, ami a pálya alakját jellemzi, és a következőképpen számítandó. Bevezetjük a

$$C = \frac{1}{N} \sum_n \mathbf{u}_n \circ \mathbf{u}_n$$

$K \times K$ méretű kovariancia mátrixot. λ_i -vel jelölve ezek sajátértékeit a komplexitás:

$$\log \Omega = - \sum_i \lambda'_i \log \lambda'_i,$$

ahol

$$\lambda'_i = \frac{\lambda_i}{\sum_j \lambda_j}$$

a sajátértékek normált változata.

Abban az esetben ha egyetlen térerősség generáló tényezőnk, azaz egyetlen független csatornánk van, akkor $\lambda_1 = 1, \lambda_i = 0, i = 2, 3, \dots, K$ és $\log \Omega = 0$. Abban az esetben, ha K darab független, nem korreláló egyenlő térgeneráló tényezőnk van C diagonális lesz és $\log \Omega = \log K$. Következésképpen $\log \Omega$ használható mint a tér szinkronizáció globális mértéke.

Az elemzésből kapott értékek vizuális tanulmányozásából kiderül, hogy a három mennyiség, Σ, Φ és $\log \Omega$ nem egymástól független, hanem az ideális gáztörvényben megjelenő nyomás, térfogat és hőmérséklethez analóg módon kapcsolatban vannak. Ennek a kapcsolatnak az értelmezése még nem tisztázott [7].

A módszert leprogramoztuk és alkalmaztuk a vadaskerti EEG jelre, de nem találtunk szignifikáns

eltérést a két csoport között valószínűleg a jelben jelenlevő zaj miatt. Ekkor vált szükségessé a nemlineáris zajszűrési módszerek alkalmazása.

2.4. Determinizmus és előrejelezhetőség

2.4.1. Előrejelezhetőség

Egy jel ami nem változik vagy periodikus annak triviális megjósolni a viselkedését amint észleltünk egy teljes ciklust a periódusból, véletlen számokat is könnyű előrejelezni, vagyis nem sokat segít a kínlódás mert a legjobb találat úgyis az átlagérték. [4] Más jelek mint például az EEG idősor valahol a kettő között van nem lehet biztosan előrejelezni de van egy bizonyos struktúrája, ami megkönnyíti a jósolást. Megpróbáljuk megjósolni a jel viselkedését majd kiszámítjuk a hibát e amire a leggyakrabban használt módszer a különbség átlagának négyzete, és eldönthetjük, hogy a jel a fentiekhez hasonlóan determinisztikus vagy sztochasztikus

$$e = \sqrt{(x_n - y_n)^2}$$

Ahol e a hiba, n a jel hossza, x_n a jel és y_n az általunk előrejelzett, létrehozott jel.

2.4.2. Egyszerű nemlineáris előrejelző algoritmus

Mint a legtöbb esetben a nemlineáris analízisnél feltételezzük, hogy a jelünk egy dinamikai rendszerből ered. Annak érdekében, hogy megjósoljuk a jövőt azaz x_n -ből x_{n+1} -et kapjunk megfigyeljük a rendszer viselkedését a múltban és megkeressük azt az állapotot x_{n_0} amelyik a legközelebb áll a jelenlegi állapothoz. Hogyha a rendszer állapota a múltban hasonló a mostani állapothoz a megfelelő dimenzióban akkor valószínű, hogy ugyanúgy fog viselkedni a rendszer mint a múltban vagyis x_{n+1} hasonlítani fog x_{n_0+1} -re. A nagy gond az, hogy nem ismerjük a bemeneti rendszer dimenzióját, és általában bemenetként egy idősorunk van, de használhatjuk a fent említett időkézés beágyazást hogy újraépítsük a jelet a megfelelő dimenzióba.

$$\mathbf{s}_n = (s_{n-(m-1)\tau}, s_{n-(m-2)\tau}, \dots, s_{n-\tau}, s_n)$$

Mint már említettük, ez az eljárás két állítható paramétert ad a rendszernek, az időkéssést τ és a dimenziót m . Ezek a megfelelő módszerekkel megbecsülhetők. Pontosabb eredményekhez vezet ha több közeli értéket vizsgálunk meg a megfelelő dimenzióban és azok átlagát feleltetjük meg az x_{n+1} lépésnek. Ennek konkrét megvalósításakor két további paraméter jelenik meg. Az a k környezet az idősorban, amelyen belül a közeli értékeket vizsgáljuk. A negyedik változót ε jelenti annak a maximális eltérésnek a mértéke, amikor két értéket még közelinek veszünk.

2.4.3. Egyszerű nemlineáris zajszűrés

A szűrésnek nem a jövő megítélése a feladata hanem, hogy a zajos jelünket helyettesítsük jobb értékekkel. Azt feltételezzük a jelről, hogy van egy determinisztikus része, amit a fenti módszerekkel ki tudunk számolni és van egy sztochasztikus része, ami a zaj. Nem egy múltbeli jelünk van hanem egy zajos jelünk ami tartalmazza egy pont szempontjából úgy a jövőt, mint a múltat és minden pontot kell helyettesíteni a neki megfelelő értékkel. A szűrésnek természetesen vannak hibái de ezek kisebbek kell legyenek, mint az eredeti jel. Tehát bemenetként van egy jelünk aminek két komponense van x_n a zaj mentes jel és η véletlen zaj.

$$s_n = x_n + \eta$$

A célunk hogy s_n -ből megkapjuk x_n -t. η -ról azt feltételezzük hogy gyengén vagy egyáltalán nem korrelál a x_n -el. A fenti előrejelző algoritmus ismeretében könnyen megkaphatjuk az eredeti jelünk, úgy hogy s_n -t helyettesítjük az előrejelzéssel, vagy ha konzervatívabbak vagyunk megteheti valamilyen lineáris kombinációja az eredeti s_n és az előrejelzett x_n jelnek.

Az algoritmus működőképes, és alkalmazható, ha ismertek a paraméterek. De a gond a módszerrel az, hogy négy számunkra ismeretlen paramétert kell meghatározni, és csak ezek után alkalmazhatjuk az algoritmust.

3. fejezet

Független komponens elemzés

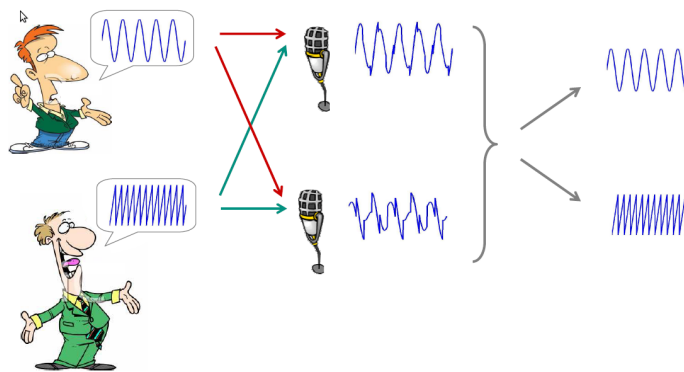
Az EEG mérései csatornákra van felosztva és minden csatorna egy elektródáról begyűjtött potenciálértékek idősorának felel meg. A csatornákat pedig egy úgynevezett referencia-pontokhoz viszonyítják, vagy a referencia-pontok átlagához. Kezdetben koponya tetején középen elhelyezett csatornához viszonyítanak, ezt egy újrareferálás követi a fül mögött elhelyezett két csatornához. Így az EEG idősorába bele kerül az elektrokardiogram (EKG) idősora is az ütőér közelsége miatt, ami nagymértékben megnehezíti az elemzést. Többek között erre a problémára is alkalmazható a független komponens analízis (Independent Component Analysis, ICA).

3.1. Mi a független komponens analízis?

Az ICA egy viszonylag új jelfeldolgozási módszer, melynek segítségével sokdimenziós adathalmazok könnyebben kezelhetők. A nemzetközi tudóstársadalom csak a 90-es évek közepén figyelt fel erre az új módszerre, ekkorra jöttek rá, hogy feladatok széles körére alkalmazható, általános algoritmusról van szó. A módszerre a legjobb példa, az úgynevezett koktél-parti probléma (3.1 ábra)[8].

Képzeljük el, hogy egy teremben egyszerre két ember beszédét vesszük fel két különböző helyen elhelyezett mikrofonnal. A mikrofonok ezáltal két kevert jelet fognak regisztrálni, melyek amplitúdóit jelölhetünk: $x_1(t)$, illetve $x_2(t)$, ahol x a megfelelő amplitúdó a t . pillanatban. A két jelet felfoghatjuk, mint a személyek által kibocsátott jelek súlyozott összege. Tehát:

$$x_1(t) = a_{11}s_1 + a_{12}s_2 \tag{3.1}$$



3.1. ábra. Koktél-parti probléma

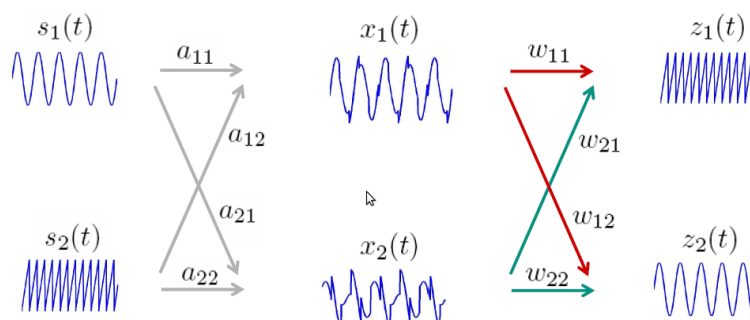
$$x_2(t) = a_{21}s_1 + a_{22}s_2 \quad (3.2)$$

Az a_{11} -, a_{12} -, a_{21} - és a_{22} -vel jelölt értékek paraméterek, melyek a mikrofonok és hangforrások egymáshoz viszonyított távolságától függenek. Ez a modell viszonylag leegyszerűsített modell, lineárisnak tekintjük a jelek egymásra-tevődését, eltekintünk az időbeli esetleges késésektől és minden egyéb extra zavarótényezőtől. Mátrixegyenlettel felírva:

$$x = As \quad (3.3)$$

Felírható az inverz egyenlet:

$$z = A^{-1}x = Wx = s \quad (3.4)$$



3.2. ábra. Keverési Mátrix

A probléma, hogy sem A sem s nem ismert, de mégis kiszámíthatók, ha feltételezzük, hogy a két jel egymástól független. Természetesen mindegy, hogy a jelek emberi beszédjelek vagy az agy különböző területei által kibocsátott elektromágneses hullámok.

3.2. Az ICA tulajdonságai

A kevert és a független jelek száma lehet eltérő is, ebben az esetben az ICA legfeljebb annyi egymástól független komponenszt talál, mint amennyi kevert jel a rendelkezésére áll. Az egyszerűség kedvéért feltételezzük, hogy ugyanannyi kevert jel van, mint amennyi független komponens, valamint a keverési mátrix is időben állandó. A módszer nem törődik a jelek amplitúdóival, hiszen a függetlenség egy állandóval történő szorzás után is megmarad. A módszer másik hiányossága talán abban van, hogy nem számít s jelek sorrendje, ez megnehezíti a dolgokat, ha az ICA-t zajszűrésre szeretnénk használni.

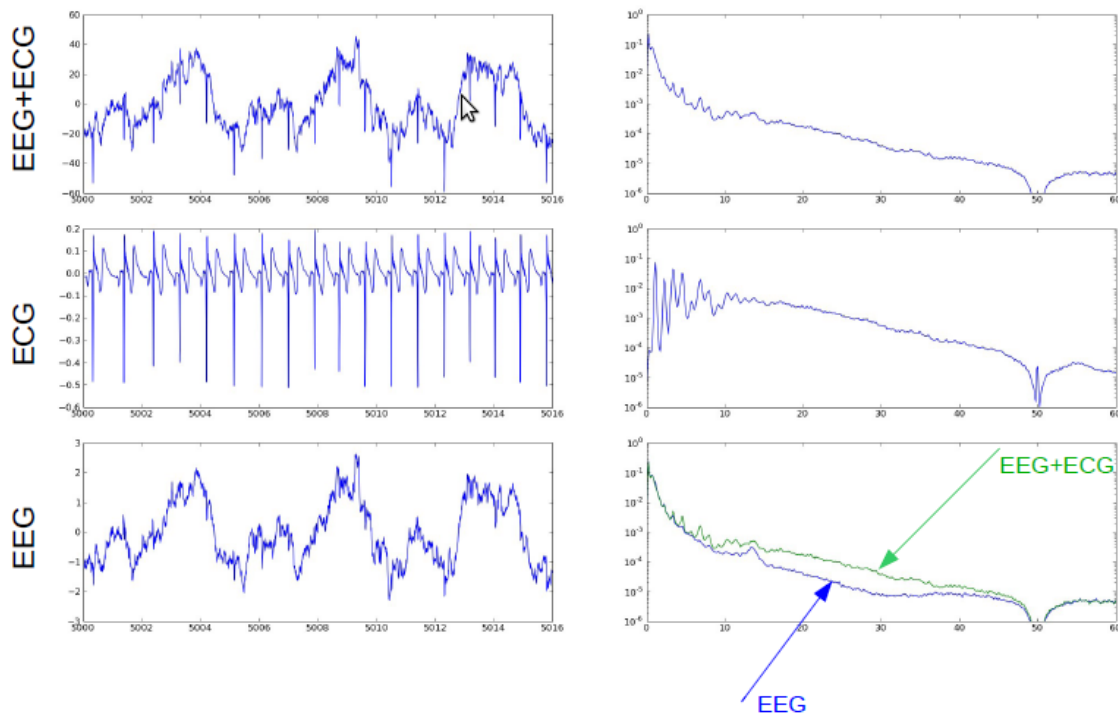
A maximálisan független komponensek meghatározásához valahogyan mérni kell a függetlenséget, ha azt feltételezzük, hogy az eredeti jelek erősen nem normális (Gauss) eloszlású jelek voltak, akkor elég csak ezt keresni a jelekben, viszont ez egy újabb feltételt von maga után, még pedig egy tiltást, mely szerint a szétválasztásra kerülő jelek nem lehetnek normális eloszlásúak, ez legfeljebb egyetlen idősor esetében engedhető meg. A nem Gauss eloszlás keresésén alapul a zajszűrésre alkalmazhatósága, ugyanis több véletlenszerű jel egymásra-tevődése tart a normális eloszláshoz. Így több módszer is adódik a függetlenség mérésére közös információk minimalizálása, a kurtózis maximalizálása vagy a negentrópia maximalizálása. A negentrópia tulajdonképpen egy normális eloszlású jel entrópiájától való eltérést fejezi ki, megmutatja a valószínűségi változó információtartalmát, minél inkább véletlenszerű, strukturálatlan a változó annál nagyobb az entrópiája. A negentrópia mindig pozitív és csak akkor nulla, ha Gauss-eloszlású változókról van szó, így a negentrópia maximalizálásával megkaphatók a leginkább nem Gauss-eloszlású változók. Egy másik mérőszám is származtatható az entrópiából, mégpedig a közös információ. A közös információ nulla akkor, ha a változók statisztikailag függetlenek, és pozitív egyébként. Azaz a közös információ minimalizálásával a leginkább független változók kaphatók meg. Szinte az összes módszer egy véletlen W inverz keverési mátrixból indul ki, majd ezt módosítja, javítja valamely fent említett jellemzőt optimalizálva.

3.3. Az ICA szerepe az alvás-elektrofiziológiában

A mi esetünkben a források szétválasztása nem teljesen vakon történik, az EKG és EEG nem teljesen egymásra-tevődött: az EEG-ben benne van az EKG de ez fordítva nem igaz, vagyis itt az ICA-ra nézve akkor jó a szűrés, amikor az EKG változatlan marad. Az ICA kiválóan alkalmazható szűrésre,

amikor a zaj mint független komponens jelforrás jelenik meg. Az EKG jel periodikus, míg az EEG jel kaotikus rezgéseknek tűnik, nem kétséges, hogy egymástól független jelekről van szó és alkalmazható lesz rajtuk az ICA. Az alábbi ábra egy mondhatni kitűnő szűrés eredménye, az ábrán látható a jel és a spektruma szűrés előtt illetve szűrés után. A szűrést a Python MDP csomagjában megtalálható FastICA[10] segítségével végeztük.

Az 3.3 ábrán megfigyelhetők a korábban tárgyalt tulajdonságok, látható, hogy a szűrés után megváltozott a referencia amplitúdója, viszont szinte tökéletesre sikerült a függetlenítés, a spektrumon megfigyelhető, hogy az elején simítás történt azon a szakaszon, ahol az EKG leginkább benne volt. Nem szabad megfélekedezni a megváltozott amplitúdóról, ezért át kell skálázni a visszaállított jelet, hogy a megengedett tartományba essen. Az amplitúdó korrekciójára a legkisebb négyzetek módszerét (3.5 egyenlet) alkalmazzuk.

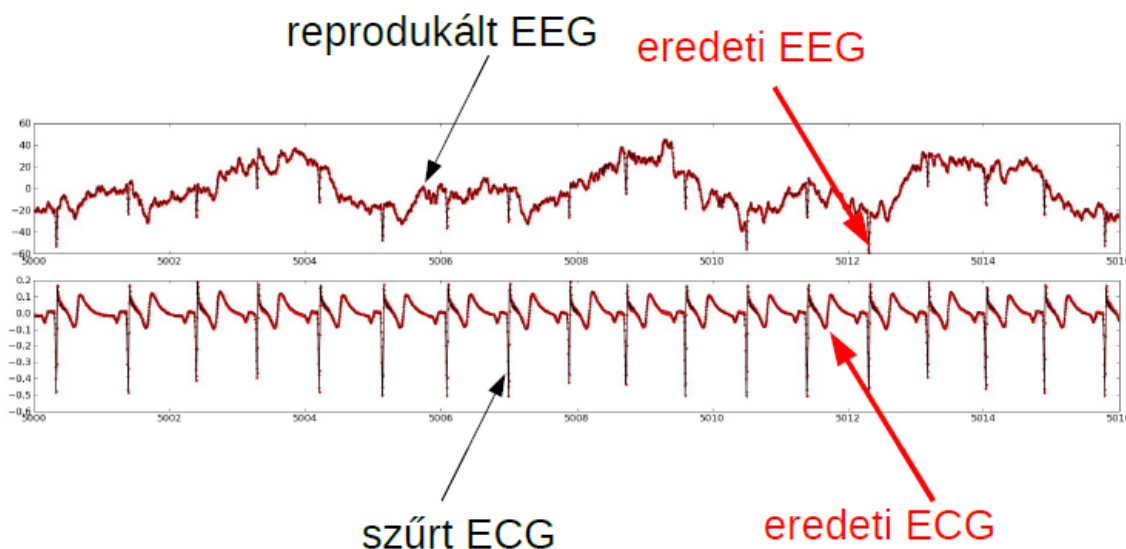


3.3. ábra. EKG szűrés az EEG referenciából, bal oldalon sorban fentről zajos EEG, EKG és szűrt EEG megváltozott amplitúdóval, míg jobb oldalon a zajos EEG, EKG majd a zajos és szűrt EEG spektrumának összehasonlítása

$$F(a, b) = \sum_{i=1}^N (z_i - ax_i - by_i)^2 = \min \quad (3.5)$$

Keressük az a és b értéket, amely segítségével x és y jelet lineárisan kombinálva minimális lesz a „távolság” z -től. Szemléltetésül, esetünkben a z lenne a zajos EEG, mely egy tiszta EEG és tiszta

EKG (feltételezés szerint) valamilyen a és b értékkel alkotott lineáris kombinációja. A két tiszta jel lenne x és y , ezeket nem ismerjük, csak az EKG-t. A módszer akkor helyes, ha reprodukálni tudjuk a zajos EEG-t, és ez sikerült is (3.4 ábra).



3.4. ábra. A reprodukált EEG a legkisebb négyzetek módszerével kapott számértékek, illetve az ICA szűréséből megmaradt jelek kombinációjából származik, és közel megegyezik a mérésből származó EEG-vel

Meg kell még említeni, hogy létezik az EEG Rapid Eye Movement stádiumába a szemmozgás által keltett zavar. Ezért elsősorban itt érdemes EOG műtermékmentesíteni. Azonban itt felmerül egy kis gond: az EOG és EKG jelek közt egy időbeli eltolódás van, ezért leghamarabb EKG mentesíteni kell az EOG-t. Természetesen ez az eltolódás befolyásolja az ICA hatékonyságát [9], ezért szűrés előtt meg kell határozni az eltolás mértékét. Az algoritmus viszonylag egyszerű, a legkisebb négyzetek módszerét használjuk a hatékonyság vizsgálatára, azt feltételezzük, hogy az EKG lineárisan eltolva tevődik rá az EOG-re, és amennyiben ez így van, akkor lesz eredményes a szűrés, ha visszkapjuk szűrés után az eredeti EKG jelet. Ugyanis ha el vannak tolódva egymástól a jelek az ICA az EKG-t elrontja, EOG-vel „fertőzi”.

4. fejezet

Hálózatelemzési módszerek

A hálózatok modern elmélete, amely a kisvilág valamint a skálafüggetlen hálózatok felfedezésével indult fejlődésnek, az egyik legújabb megközelítés a komplex hálózatok tanulmányozására. Mielőtt azonban ezen elmélet az idegtudományban való konkrét alkalmazásaira térnénk, lássuk előbb a hálózatok és ezek néhány fontosabb tulajdonságainak leírását.

4.1. A modern hálózatelmélet alapjai

A gráfok a hálózatok egy absztrakt értelmezése, melyek csúcsokból és élekből állnak. Két csúcs közötti él jelenléte valamilyen kölcsönhatásra, kapcsolatra utal az illető két csúcs között. Az egy csúcshoz tartozó élek száma megadja az illető csúcs fokszámát, annak a valószínűségét, hogy egy véletlenszerűen kiválasztott csúcsnak a fokszáma k legyen a fokszámeloszlás adja meg.

Beszélhetünk súlyozatlan és súlyozott gráfokról. Súlyozott gráfok esetében két csúcs közötti él vagy létezik, vagy nem, súlyozott gráfok esetében minden élhez hozzárendelünk egy értéket, súlyt, amely utalhat az illető él által összekötött csúcsok közötti kapcsolat erősségére, vagy a csúcsok közötti távolságra.

A súlyozatlan gráfok lokális és globális szerkezetének jellemzésére leggyakrabban használt két mennyiség a klaszterizációs együttható C és az átlagos úthossz L . A k_i fokszámú i . csúcs klaszterizációs együtthatóját C_i általában az i . csúcs szomszédai között létező élek számának és az i . csúcs szomszédai között lehetséges összes él számának arányaként értelmezzük (egy csúcsot, akkor tekintünk az i . csúcs szomszédjának, ha van közöttük él).

A gráf klaszterizációs együtthatója a C_i klaszterizációs együtthatók átlagaként adható meg.

A klaszterizációs együttható a gráf lokális szerkezetéről ad információt, valamint a véletlenszerű hibákkal szembeni rugalmasság mértékeként is értelmezték (ha az i . csúcs elvesztődik a szomszédai kapcsolatban maradnak-e a gráffal).

Egy másik fontos tulajdonsága a gráfoknak az átlagos úthossz L . A súlyozatlan gráfok esetében az i . és j . csúcsok közötti távolság $d_{i,j}$ egyenlő azzal a minimális élszámmal, amelyen áthaladva az i . csúcsból a j . csúcsba jutunk. A gráfra jellemző úthossz az összes lehetséges csúcs párok közötti úthosszak átlagaként adható meg.

A gráfra jellemző átlagos úthossz egy globális tulajdonság, mely jelzi, hogy mennyire egységesített (integrált) az illető gráf. Egy másik, az úthosszhoz kapcsolódó mennyiség, amelyet számolni szoktak a gráf átlója: a $d_{i,j}$ úthosszak közül a leghosszabb.

A fokszámeloszlás, a klaszterizációs együttható és a karakterisztikus úthossz az alapmértékei a hálózatoknak. Ezen mértékek alapján négy gráf típust különböztethetünk meg: (i) rendezett vagy rácsszerű, (nagy C . és hosszú L jellemző) (ii) kisvilág, (C közel van a rendezett hálózatokéhoz, de nagyon kicsi az L , közeli a véletlen hálózatokéhoz) (iii) véletlen (C kicsi, K/N , és L rövid, $\log N / \log K$, ahol N a csúcsok száma, K az átlagos fokszám) (iv) skálafüggetlen (az L nagyon kicsi lehet, $\log \log N$ nagyságrendű, és C is kisebb, mint a kisvilág hálózatoknál) hálózatok.

Az előzőekben részletesen tárgyalt tulajdonságokon kívül más, a hálózatok jellemzésére szolgáló tulajdonságokat is bevezettek, ilyenek a: (i) gráf sűrűség, azt mutatja meg, hogy az összes lehetséges kapcsolatból mennyi található meg a valóságban; (ii) közelség, megadja hogy egy csúcsot milyen könnyen lehet elérni a többi csúcsból, (iii) kohézió, két csúcs esetében megadja, hogy legkevesebb hány élet kell kivenni, ahhoz, hogy az illető két csúcsot elválasszuk egymástól, az egész gráfra értelmezve az összes lehetséges csúcs párra számolt kohézióértékek közül a legkisebbet adja, (iv) köztesség (betweenness), amely egy adott i csúcsra, a j és k közötti, i -t is tartalmazó minimális úthossz, és a j és k közötti minimális úthossz arányának az összege az összes lehetséges j, k párosításra

Az eddig említett tulajdonságok a súlyozatlan gráfokra vonatkoztak, azonban nagyon sok esetben a súlyozott gráfok egy sokkal precízebb modellt nyújtanak a valós hálózatok leírására. A súlyozott gráfok súlyozatlan gráfokká alakítása számos hátránnyal bír: sok, a súlyozott gráfok esetében létező információt nem veszünk figyelembe, ha a küszöbérték túl magas néhány csúcs elszigeteltté válhat, a küszöbérték megválasztása tetszőleges. A gráfok és ezek tulajdonságainak részletesebb leírását lásd a [11] cikkben.

| Faj | L | L_{rand} | C | C_{rand} |
|------------------------------|------|-------------------|------|-------------------|
| <i>C. elegans</i> | 2.65 | 2.25 | 0.28 | 0.05 |
| Makákó majom (vizuális) | 1.69 | 1.65 | 0.59 | 0.32 |
| Makákó majom (somatosensory) | 1.77 | 1.72 | 0.57 | 0.31 |
| Makákó (teljes agykéreg) | 2.18 | 1.95 | 0.49 | 0.16 |
| Macska (teljes agykéreg) | 1.79 | 1.67 | 0.6 | 0.3 |

4.1. táblázat. Mind az öt esetben látszik, hogy az idegi hálózat jellemző úthossza (L) közel akkora, mint az ugyanannyi csúcsot tartalmazó és ugyanolyan átlagos fokszámmal rendelkező véletlen hálózat úthossza (L_{rand}), valamint az idegi hálózat klaszterizációs együtthatója (C) jóval nagyobb, mint a véletlen hálózat esetén kapott érték (C_{rand}), amely megerősíti azt a feltételezést, hogy az idegi hálózatok kisvilág szerkezetűek.

4.2. Idegtudományi alkalmazások

Lago-Fernandez a három különböző alap hálózat-szerkezeti típus az idegsejt-hálózatok koherens oszcillációira gyakorolt hatását tanulmányozta [12]. A véletlen hálózatok esetében gyors volt a rendszerválasz, azonban nem tudtak koherens oszcillációkat kelteni. A rendezett szerkezetű hálózatok koherens oszcillációkat mutattak, de nem volt gyors a jel feldolgozás. A kisvilág hálózatokra azonban jellemző volt úgy a gyors rendszerválasz, mint a koherens oszcillációk keltése, azt sugallva ezzel, hogy ez a szerkezeti típus a legmegfelelőbb az ideghálózatok információ feldolgozásának modellezésére.

Watts and Strogatz a *Caenorhabditis elegans* idegrendszerének - amely az egyetlen teljesen feltérképezett idegi hálózat - anatómiai kapcsoltságát tanulmányozták [13]. Ez az idegi hálózat egy $N = 282$ csúcsot tartalmazó és $k = 14$ átlagos fokszámú gráffal modellezhető. A neuronok összekötöttek tekinthetők, ha van köztük szinapszis vagy réses kapcsolat. A gráfot megvizsgálva, kimutatták, hogy kisvilág hálózatra jellemző a klaszterizációs együtthatója és az átlagos úthossza. Ez volt az első bizonyíték az idegrendszer kisvilág hálózati szerkezetére. Hasonló következtetések vonhatóak le más tanulmányokból is [14] [15] (lásd 4.2 táblázatban):

Az első többé kevésbé direkt bizonyítása az anatómiai kapcsoltság kisvilág szerű hálózati szerkezetének egy MRI jel vizsgálata volt: 124 egészséges személyt vizsgáltak, két zónát kapcsolatban levőnek tekintettek, ha statisztikailag jelentős korreláció volt közöttük.

Feltevődik a kérdés, hogy miért is van az agyi hálózatoknak kisvilág szerkezetük. Feltételezzük, hogy az agy szerkezete úgy alakult ki, hogy maximalizálja a komplexitást, míg a „költségeket” minimalizálja (lehető legkönnyebben, legrövidebb úton tudjunk eljutni bármely csúcsból bármely másik

csúcsba).

Sporns és kollégái abból a feltételezésből indultak ki, hogy az agy két egymással ellentétes követelménynek kell eleget tennie: (i) szegregáció, azaz egy lokális specializáció sajátos feladatokra, és (ii) integráció, egy globális szintű egyesítése az információknak. Tanulmányaik során arra a következtetésre jutottak, hogy a komplexitás szempontjából - az optimális egyensúly a szegregáció és integráció között - a legmegfelelőbb hálózat típus a kisvilág hálózat. Különböző más tanulmányokból az is kiderült, hogy a költségek minimalizálására is a kisvilág hálózat felel meg a leginkább [16].

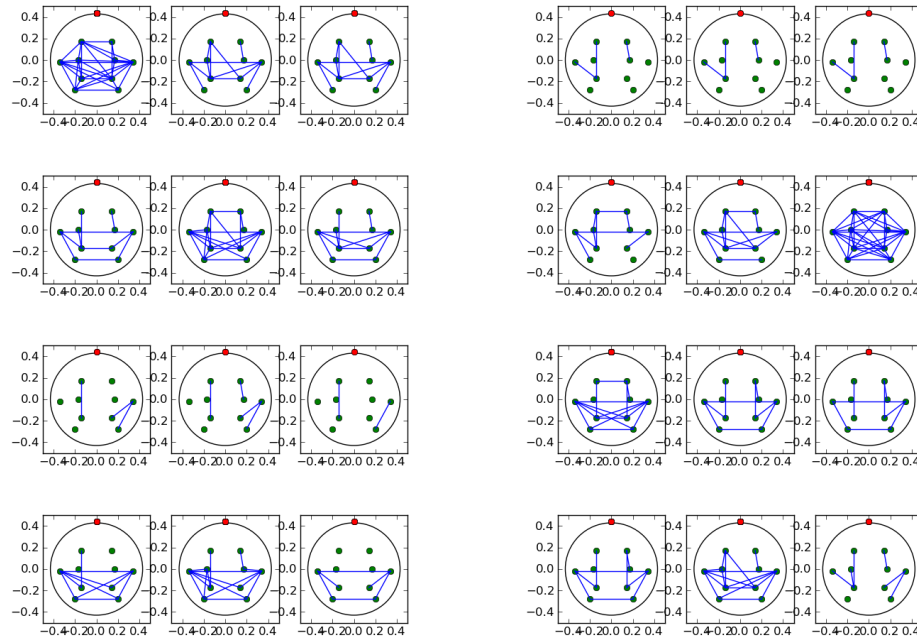
Egy másik jelentős hálózatelméleti alkalmazás a Stam és kollégái által végzett analízis az Alzheimer kóros betegek EEG jelének [17]. Az egészséges alanyok és az Alzheimer kóros alanyok eredményeit összehasonlítva azt vették észre, hogy a beteg alanyok EEG jeléből kapott gráfok L úthossza lényegesen hosszabb volt mint a kontroll egészséges alanyok esetén, míg a C klaszteriizációs együttható változása nem volt szignifikáns. Tehát az Alzheimer kóros betegek esetén a kisvilág szerkezet kevésbé jellemző az agyi hálózatuk szerkezetére mint az egészséges alanyok esetén.

4.3. Funkcionális hálózatok az alváskutatásban

A korábban említett, és egyéb, a hálózatok tulajdonságait vizsgáló algoritmusokat alkalmaztunk EEG jelre. Az általunk feldolgozott EEG jel két lényegesen különböző csoporttól származott: egy kísérleti autista csoporttól és egy kontroll egészséges csoporttól. A mintavételezés egész éjszakai alvás közben történt, amely három alvásciklusra volt osztva, 256 Hz-es frekvencián. Elsődleges célunk az volt, hogy valamilyen szignifikáns eltérést mutassunk ki a kísérleti és kontroll csoport EEG jeléből számolt valamely mennyiség(ek) között, illetve magyarázatot adni az észlelt különbségre.

Első lépésben az EEG adatoknak egy hálózati ábrázolását kell megoldjuk. Tudjuk, hogy minden alany fejére 10 darab elektródát helyeztek, így tehát 10 csatornánk van, ezek lesznek a gráf csúcsai. Ahhoz, hogy az éleket is meghatározzuk kiszámoltuk a csatornák közötti koherenciát az összes lehetséges párosítás esetén. Ezek után, ha súlyozott gráffal akarunk dolgozni, azt mondjuk, hogy az i . és j . csúcsok közötti él súlya az i . és j . csatornák közötti koherencia átlagértékével egyenlő; vagy tekinthetjük súlyozatlannak a gráfot, ebben az esetben meghatározunk egy küszöb-koherenciaértéket, és ha a koherencia átlagérték ez az érték alatt van, akkor az élet 0-nak vesszük, azaz nincs él, ellenkező esetben 1-nek, azaz létezőnek vesszük. Az így keletkezett gráfok (4.1) ábrán megtekinthetők.

Tekintsük először a súlyozatlan gráfok esetét. Kiszámoltuk a koherencia átlagértéket minden alany



4.1. ábra. A 0.4 küszöb-koherenciaérték esetén keletkezett gráfok, a jobb oldali három oszlop a kísérleti alanyok három alvásciklusának felel meg, a bal oldali a kontroll alanyok három alvásciklusának felel meg, négy kísérleti és négy kontroll alanyra ábrázolva

mindhárom alvásciklusára a következő négy frekvenciatartományban: 0.5Hz-1.25Hz, 0.5Hz-4Hz, 1.25Hz-4Hz, valamint 10Hz-15Hz. Most következne a küszöbérték lerögzítése. Mivel azonban nem tudhatjuk, hogy a küszöbérték milyen megválasztásával kapunk szignifikáns eltérést, nem rögzítettük, hanem végimentünk vele a $[0,0.8]$ intervallumon 0.02-es léptékkel, és minden egyes küszöbértékre vizsgáltuk a négy különböző frekvenciatartományban, alvásciklusonként a kísérleti és kontroll csoport EEG jeléből kapott gráf valamely tulajdonságát. (Mivel a koherencia 0 és 1 között bármilyen értéket felvehet, logikusnak tűnne, hogy a küszöbértékkel is a $[0,1]$ intervallumon menjünk végig, azonban észrevettük, hogy a csúcsok közti koherenciára 0.8-nál nagyobb értéket egyik frekvenciatartományban sem kaptunk, így fölösleges annál nagyobb küszöbértékek esetét vizsgálni).

A vizsgált tulajdonságok a következők voltak: klaszterizációs együttható, köztesség, átlagos úthossz, átlagos foksám, gráf átmérő, gráf sűrűség, közelség (hogy egy, a gráfra jellemző tulajdonságot kapjunk, az összes csúcsra vett átlagértékkel dolgoztunk) kohézió.

Minden tulajdonság esetén összehasonlítottuk a kísérleti és kontroll csoport eredményeit: egy

szignifikancia tesztet végeztünk rajtuk, amely a két különböző minta átlagát hasonlítja össze, és megadja annak a valószínűségét, hogy a két minta ugyanattól a populációtól származik, tehát annak a valószínűségét, hogy az átlagok különbségének oka csak statisztikai fluktuáció, és nem utal semmilyen, a minták közötti lényeges különbségre.

Ezen tulajdonságok vizsgálatával egyik frekvenciatartományban sem, egyetlen alvásciklus és semmilyen küszöbérték esetén nem kaptunk szignifikáns különbséget a kísérleti és kontroll csoport eredményei között (szignifikánsnak tekinthető a különbség, ha az egy populációhoz tartozás valószínűsége kisebb mint 5%).

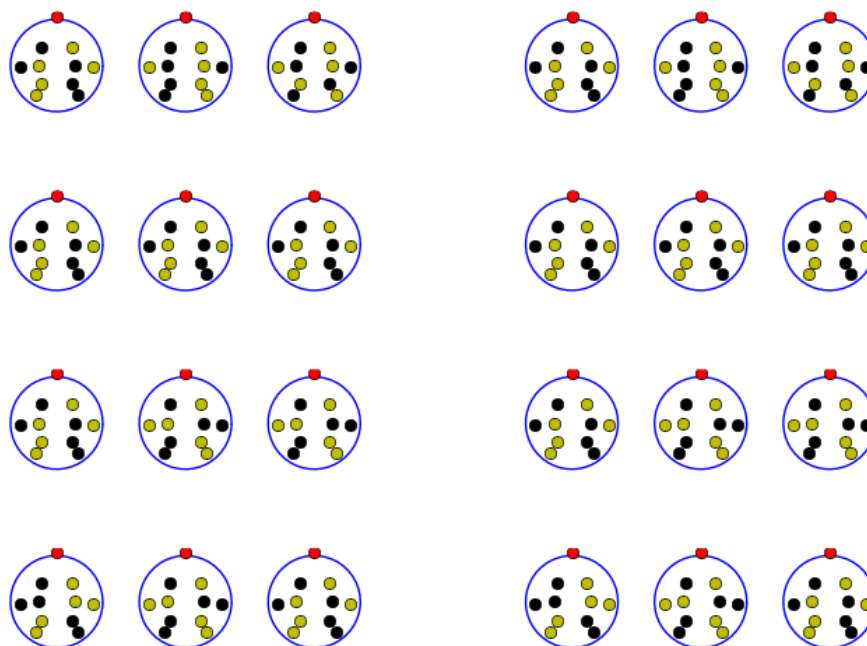
Mivel ez nem vezetett lényeges eredményre, a következőekben azt vizsgáltuk, hogy hogyan lehetne a keletkezett gráfokat minimális illetve maximális költséggel két egyenlő számú csúcsot tartalmazó csoportra osztani. Ebben az esetben a gráfokat a fentebb leírt módon súlyozottnak tekintettük. Az elosztás költsége egyenlő azon élek súlyainak összegével, amelyeket fel kell számoljunk a csúcspontok két különböző csoportba osztásakor (tehát az összes olyan él súlyának az összege, amely egyik csoportból bármely csúcsot a másik csoport bármely csúcsával összeköt). A kapott eredményt ábráztuk minden frekvenciatartományban(4.2).

Ránézésre nem sok különbség észlelhető a kísérleti és kontroll csoportra kapott eredmények között, így valamivel jellemezni kellett a kettéosztás utáni szerkezetet, hogy a kísérleti és kontroll csoport közötti különbséget vizsgálni tudjuk. Egy dipólus szerű mennyiséget számoltunk: mintha a sárgák pozitív, a feketék meg negatív töltések lennének. A dipólus komponensei a (4.1) képlettel a következőképpen adhatók meg:

$$D_i = \sum_{\alpha=1}^{10} a^\alpha n_i^\alpha, \quad (4.1)$$

ahol, $a^\alpha \in \{-1, 1\}$, attól függően, hogy melyik csoportban van (piros vagy kék), és n_i^α az illető csúcs helyzetvektorának komponensei. Ismerve az elektródák helyének koordinátáit (ezek adottak két dimenzióban, poláris koordináták formájában), valamint minden alanyra az elosztást, ez a dipólus szerű mennyiség könnyen kiszámolható.

A tulajdonságok vizsgálatánál tárgyaltakhoz hasonlóan ebben az esetben is vizsgáltuk, hogy van-e szignifikáns különbség a dipólus nagyságokban a kísérleti és kontroll csoport eredményei között a négy különböző frekvenciatartományban, a három alvásciklus esetén, minimum és maximum költséggel történő elosztáskor. Az eredmények itt biztatóbbak voltak (főleg maximális költséggel való kettéosztások esetén gyakori volt a 0.08-0.10 közötti valószínűség, arra hogy azonos populációtól



4.2. ábra. kettéosztás eredménye, 0.5Hz-1.25Hz frekvenciatartományra, minimális költséggel történő elosztás esetén; sárgával az egyik csoporthoz, feketével a másik csoporthoz tartozó csúcsok; jobb oldali három oszlop a kísérleti, bal oldali három oszlop kontroll alanyok, alvásciklusonként, négy kísérleti és négy kontroll alany van ábrázolva

származik a mérési eredmény) de összességében ez sem tekinthető szignifikáns eredménynek.

Megvizsgáltuk azt is, hogy a különböző lehetséges csúcsok (csatornák) kísérleti és kontroll csoport esetében a kettéosztás után hányszor voltak azonos csoportban és hányszor különböző csoportban. A maximális költséggel történő elosztást vizsgálva észrevettük, hogy mindhárom alvásciklus, és mind a négy frekvenciatartomány esetén az F3 és O2 csatornák, valamint a C4 és O2 a kísérleti alanyoknál mindig külön csoportban voltak, a kontroll alanyoknál viszont szinte mindig ugyanabban a csoportban. A minimális költséggel történő elosztás esetében csak a 10Hz-15Hz frekvenciatartományban volt időben állandónak tűnő különbség észlelhető: a C3 és P4, C3 és T4 valamint a C4 és P3 csatornák a kísérleti alanyoknál ugyanannyiszor voltak különböző, mint azonos csoportban, míg a kontroll alanyoknál szinte mindig azonos csoportban voltak (a csatornák helyét és elnevezését lásd 1.2a ábrán).

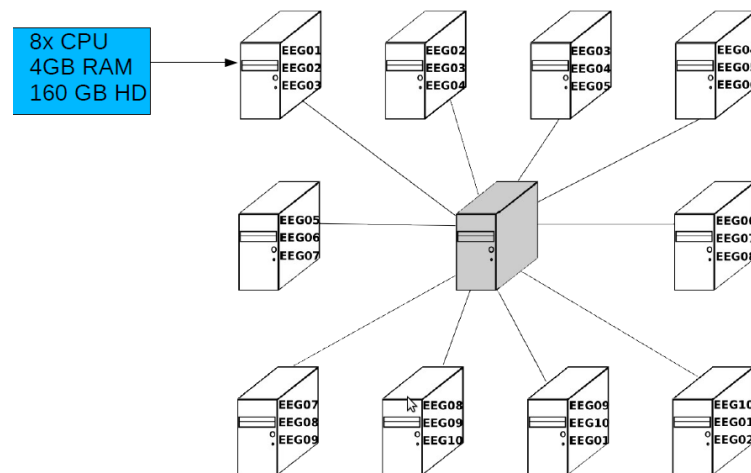
Azonban egy egészen más megközelítésből nézve a dolgokat egy érdekes dolog figyelhető meg.

Függetlenül attól, hogy kísérleti vagy kontroll csoportról van szó, egy személy három különböző alvásciklusában a kettéosztás eredménye szinte mindig teljesen ugyanaz, vagy legalább nagyon hasonló. Más szóval, sokkal nagyobb a szórás az emberek között, mint egy személyen belül időben. Ez egy úgynevezett ujjenyomat viselkedés, ami személyre jellemző.

Ennek, bizonyítás érdekében a három alvásciklust felosztottuk félórás szakaszokra, ezekre is koherenciát számoltunk, majd ugyanúgy, mint ahogy előzőleg a személyekre, most a félórás szakaszokra számoltunk dipólust minden személy esetén. Ezek után a személyek menti egész éjszakára számolt dipólus nagyságok szórását hasonlítottuk össze az idő menti dipólus nagyságok szórásával (mivel idő mentén minden személyre külön szórást számoltunk, az idő menti szórás ezen személyenkénti szórások átlagaként adható meg), eredményül azt kaptuk, hogy a 10Hz-15Hz frekvenciatartományban az idő menti szórás hozzávetőlegesen egy nagyságrenddel kisebb, mint a személyek dipólus nagyságának szórása, és a többi frekvenciatartományokban is legalább két vagy háromszorosa a személyek menti szórás az idő menti szórásnál, mely eredmény alátámasztja a feltételezésünket.

5. fejezet

Optimalizált számítógépes infrastruktúra



5.1. ábra. Adatok szétszttása a klaszteren

Nagy mennyiségű adat elemzéséhez és feldolgozásához nagy teljesítményű számítógépekre van szükség, manapság túl nagy mennyiségű adathalmazokkal dolgoznak a tudományok különböző területén, egy számítógép nem is elég a feldolgozásukhoz. Erre lettek kifejlesztve a klaszterek, azaz több számítógépből álló hálózati rendszerek. A módszerünk lényege, hogy a különböző személyek teljes éjszakás jelei különböző ágon lett elhelyezve (5.2 ábra) és az ágak mindenkire a gyökéren definiált algoritmust végzik el. Ehhez szükség van magára az algoritmusra, egy konfigurációs fájlra¹ és a futtatóra.

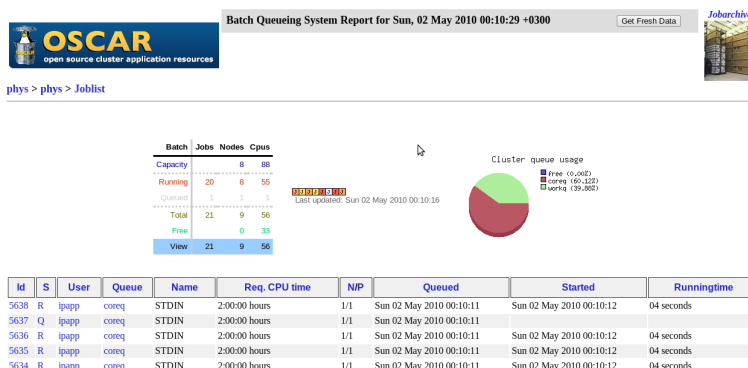
Példa konfigurációs állományra:

```
jobname = spectrumtest
```

¹ Konfigurációs állománnyal megadjuk a futtatónak, hogy melyik állomány milyen ágon van

```
[jobs]
[01]]
command = python spectrum2.py 256
infile = /eeg/vadaskert08/AcsB/1cic.dat, \
        /eeg/vadaskert08/AcsB/2cic.dat, \
        /eeg/vadaskert08/AcsB/3cic.dat
outfile = AcsB_1cic, AcsB_2cic, AcsB_3cic
node = oscar05.cl
```

A konfigurációs fájlban tartalmaznia kell a parancsot, a mintavételi frekvenciát bemenő, valamint ki-menő állományok nevét és, hogy melyik ágon fog futni. A futtató a megadott paraméterekkel fog dol-gozni, a megadott állományokat csak a megadott ágakon fogja futtatni. A futtató bemeneti paraméter-ként kapja a konfigurációs állományt, míg az algoritmus bemeneti paraméter-ként. Ezzel a módszerrel éltünk a spektrum számolások futtatásánál.



5.2. ábra. A munkafolyamatok követése

6. fejezet

Következtetés és kitekintés

A dolgozatban egy sor olyan matematikai, numerikus és számítástechnikai módszert mutattunk be, melyek az elektrofiziológiai adatok korszerű elemzéséhez elengedhetetlenek. Ezek között olyanok is vannak, melyek először kerülnek alkalmazásra ezen a területen. Ezeket valós adatokon fejlesztettük ki sok esetben újszerű tulajdonságokra bukkanva a jelben. Az eredmények közül kiemelendő elsősorban a hálózati módszerekkel felfedezett ujjlenyomat tulajdonsága a koherenciának. Megmutattuk ugyanakkor, hogy miként lehet egy sokmagos számítógépes rendszert optimális módon az EEG adatelemzés szolgálatába állítani.

A továbbiakban a kifejlesztett módszerek egy részére statisztikai elemzést kell végeznünk, és feltételezzük, hogy a nemlineáris szűrés eredményeképpen szignifikáns eredmények jelennek meg ott is, ahol korábban csak tendenciák voltak észlelhetők.

Irodalomjegyzék

- [1] J. L. Gischer, The equational theory of pomsets. *Theoret. Comput. Sci.*, **61**(1988), 199–224.
- [2] J.-E. Pin, *Varieties of Formal Languages*, Plenum Publishing Corp., New York, 1986.
- [3] C.J. Stam, Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field, *Clin. Neurophys.* (2005)
- [4] Holger Kantz, Thomas Schreiber, *Nonlinear Time Series Analysis Max Planck Inst.* (2004)
- [5] Béla Weiss, Zsófia Clemens, Róbert Boldizs, Zsuzsanna Vágó, Péter Halász, Statio-temporal analysis of monofractal and multifractal properties of the human sleep EEG *Faculty of I.T. Pazmany P.* (2009)
- [6] Jong-Min Lee, Dae-Jin Kim, In-Young Kim, Kwang Suk Park, Sun I. Kim, Nonlinear-analysis of human sleep EEG using detrended fluctuation analysis *Dep. Biomedical E.Hanyang U.* (2004)
- [7] Jiri Wackermann Towards a quantitative characterisation of functional states of the brain: from the non-linear methodology to the global linear description *I.Grenzgebiete der Psych.,Germany* (1999)
- [8] Aapo Hyvärinen, Erkki Oja: „Independent Component Analysis: Algorithms and Applications” *Neural Networks April*, 1999
- [9] Shifted Independent Component Analysis Morten Morup, Kristoffer H. Madsen, and Lars K. Hansen *Lyngby Denmark*
- [10] Aapo Hyvärinen: „Fast and Robust Fixed-Point Algorithms for Independent Component Analysis” *Helsinki University of Technology, Laboratory of Computer and Information science, Januar 22*, 1999

- [11] Reka Albert and Albert-Laszlo Barabasi: Statistical mechanics of complex networks. *Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556* (2002)
- [12] Lago-Fernandez LF, Huerta R, Corbacho F, Siguenza JA: Fast response and temporal coherent oscillations in small-world networks. *Phys Rev Lett* 2000, 84:2758-2761.
- [13] Watts DJ, Strogatz SH: Collective dynamics of "small-world" networks. *Nature* 1998, 393:440-442. .
- [14] Cornelis J Stam and Jaap Reijneveld: Graph theoretical analysis of complex networks in the brain *Nonlinear Biomedical Physics* 2007
- [15] Danielle Smith Bassett and Ed Bullmore: Small-World Brain Networks *The Neuroscientist* Volume 12, Number 6, 2006
- [16] Sporns O, Tononi G, Edelman GM: Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cereb Cortex* 2000, 10:127-141.
- [17] C.J.Stam, B. F. Jones, G. Nolte, M. Breakspear and PH. Scheltens: Small-World Networks and Functional Connectivity in Alzheimer's Disease *Cerebral Cortex* January 2007;17:92-99