

Posudek na dizertační práci: **Exploring Brain Network Connectivity Through Hemodynamic Modeling**

Autor práce: **Ing. Martin Havlíček**

Předložená práce se zabývá aplikací metod Bayesovské statistiky na problém identifikace struktury propojení mozkových oblastí. Práce je přehledně rozdělena do tří hlavních částí, které jsou v kapitolách 2 až 3. V kapitolách 2 a 4 si autor připravuje teoretické výsledky, které aplikuje na zvolený problém v kapitole 4. Nejprve se budu zabývat postupně každou hlavní kapitolou.

Kapitola číslo 2 je věnována problému sdružené identifikace stavů a parametrů nelineárního dynamického systému. V této kapitole autor představuje standardní techniky využívané pro zmíněný účel a jejich posouzení s ohledem na studovaný problém. Jako nejvhodnější je zvolena technika Cubature Kalman filtering (CKF). Tuto techniku rozšiřuje s využitím známých výsledků na úlohu hlazení (smoothing) a na odhadování parametrů systémů se stavovým modelem ve spojitém čase. Další použitou technikou je Variálně Bayesovské odhadování variancí stavového šumu a šumu pozorování. Z těchto ingrediencí je složen celkový algoritmus použitý k odhadu hemodynamického modelu.

Přehled uvedených technik je téměř vyčerpávající a jejich posouzení vesměs správné. Vybrané techniky se zdají být vhodné pro řešený problém a v současnosti by bylo obtížné nalézt modernější, či perspektivnější alternativu. Spojení těchto algoritmů do jediného celku je výborný výsledek a je možné jej považovat za přínos práce. Ke kapitole bych měl následující dotazy:

- CKF je prezentován na straně 26 jako jasná volba na základě bonmotu. Gausovská aproximace aposteriorna je skutečně jednoduchá, ale bývá nedostatečně přesná pokud je aposteriorno multimodální a nebo existuje-li v prostoru parametrů varieta na níž je věrohodnostní funkce konstantní. Jaké indície naznačují, že ani jeden problém nenastane?
- Na začátku části 2.4. uvádíte jako cíl práce on-line odhad parametrů. Je to z důvodu předpokladu časové proměnnosti parametrů nebo je to z jiného důvodu?

Za hlavní cíl si autor vytyčil určení kauzálního propojení mozkových oblastí a nástroje pro tento cíl prezentuje v kapitole číslo 3. Nejprve je představen stochastický model propojení různých oblastí mozku a poté jsou prezentovány metody Bayesovského výběru modelu.

Model propojení je prezentován krátce avšak jeho napojení na model hemodynamiky je věnována malá pozornost. Schéma na obrázku 3.2 je graficky efektní, ale mnoho nevysvětluje – například se v něm stavová veličina $\hat{z}$ nevyskytuje. Tato část by zasloužila více pozornosti, neboť je zde více možností než jen složení všech veličin do jednoho

velkého modelu. Dále se zde mluví o Kalmanově filtru, není zřejmé, zda jde o standardní Kalmanův filtr nebo o metodiku odhadu z kapitoly 2. Naopak část o Bayesovském výběru modelu jsou popsány důkladně a jejich varianty pro jednodušší výpočet také. K části mám následující otázky:

- Na straně 75 je uvedeno, že Kalmanův filtr není plně Bayesovský a není možno definovat apriorno. Proč není Kalmanův filtr Bayesovská metoda?
- Z popisu jsem nabyl dojmu, že všechny parametry jsou spojené do jediného stavového vektoru a odhadnuty sdruženě. Avšak hemodynamické modely pro jednotlivé oblasti mozku jsou podmíněně nezávislé za podmínky neurologického signálu. Tato strukturální vlastnost tedy není brána do úvahy při návrhu odhadovací metody. Bylo by výhodné tuto vlastnost využít?
- Na straně 86 uvádíte, že Kalmanův filtr není schopen odhadovat apriorní hodnoty pro ARD. Jaké rozšíření by bylo potřeba udělat aby bylo možné ARD použít?
- Při odhadování struktury modelu metodou prořezávání využíváte diagonálu matice přesnosti. To je možné interpretovat jako převod na nezávislé parametry. Je nějaký důvod proč nepoužít přímo marginalizaci (optimální převod na nezávislé parametry) a použít diagonálu kovarianční matice?

Navržené modely jsou testovány v kapitole 4 na simulovaných a reálných datech. Experimenty na simulovaných datech mají za cíl představit metodu (4.2.1), odhadnout citlivost na parametry (4.2.2), představit a otestovat spolehlivost odhadu struktury modelu na pro různé úrovně šumu (4.3.2) a doby vzorkování (4.3.3), dalších neurologické variace (4.3.5–4.3.6). Na reálných datech je metoda otestována v části (4.3.7).

V této kapitole je prezentováno široká škála experimentů a počet provedených Monte Carlo simulací je úctyhodný. Ilustrační část 4.2.1 je velmi dobře provedena a vhodně ilustruje problém. V části 4.2.2 je aplikována technika odhadu struktury modelu na hemodynamický model se šesti parametry. Závěrem je jako nejvhodnější model vybrán model pouze se dvěma parametry. Není však jasné, jestli bude pro následující experimenty použit původní plný model nebo ten nový odhadnutý. Tento experiment považuji za zásadní, avšak jeho interpretace se jeví jako sporná. Výsledek totiž závisí na dvou faktorech: (i) skutečné identifikovatelnosti modelu a (ii) přesnosti aproximativní metody SKCS. Z tohoto experimentu není možné rozlišit, který faktor má větší vliv. Obojí má značné důsledky na interpretaci ostatních výsledků. Reálná data bývají hlavním testovacím kamenem každé metody. Je s podivem, že jim byla v této práci věnována tak malá pozornost. Experiment na reálných datech je tedy spíše příslibem do budoucna, že metoda metoda je perspektivní a ověřitelná. Otázky ke kapitole:

- Jaký model hemodynamické odezvy byl použit při experimentech, 1 nebo 11? Je možné provést experiment z části 4.2.2 na reálných datech?
- Jaké jsou možnosti ověření výsledků metody na reálných datech s ostatními přístupy?

V závěru práce jsou uvedeny přínosy práce a její zhodnocení. Autor si je vědom jejích slabin a má představu jak je řešit.

Pro celkové zhodnocení práce je třeba přihlédnout k tomu, že téma práce je velice perspektivní, náročné a ambiciózní. Celá oblast výzkumu se rychle rozvíjí, není mnoho zaručených a spolehlivých tvrzení. Tím, že nové poznatky, metody a data rychle přibývají je třeba neustále reagovat na měnící se situaci. Za hlavní přínos práce považuji systematickou aplikaci Bayesovské teorie do oblasti modelování biologických procesů a zpracování dat. Představené metody nejsou voleny náhodně. Autor vychází ze spolehlivých teoretických základů a aproximace potřebné k výpočtu výsledku pečlivě vybírá s ohledem na zvolený cíl. Autor využívá nejmodernější metody a kombinuje je do harmonického celku, který se zdá být velmi vhodným nástrojem pro studium zvolené problematiky.

Výsledky prezentované v této dizertaci byly již částečně publikovány, avšak jejich publikační potenciál zdaleka nebyl vyčerpán. Autor prokázal schopnosti, které mu dovolují v práci pokračovat na profesionální úrovni a dosáhnout ještě významnějších výsledků v blízké budoucnosti. Za příkladnou považuji především autorovu práci s literaturou, je schopen si rychle osvojit nejnovější metody, zakomponovat je do obecné teorie a využít pro zvolený cíl.

Práci považuji za velmi kvalitní a doporučuji k obhajobě.

Ing. Václav Šmídl, Ph.D.