

POROVNÁNÍ DETEKTORŮ VYSOKOFREKVENČNÍCH OSCILACÍ V INTRAKRANIÁLNÍM ELEKTROENCEFALOGRAMU

T. Havel, P. Ježdík

Katedra teorie obvodů, Fakulta elektrotechnická, České vysoké učení technické v Praze

Abstrakt

V minulosti byla prokázána souvislost mezi epileptogenní zónou a oblastí v níž se vyskytují vysokofrekvenční oscilace (HFO). Aby mohla být tato skutečnost plně využita je potřeba identifikovat a lokalizovat HFO v dlouhodobém elektroencefalogramu. Vizualní hodnocení je časově příliš náročné, tudíž je potřeba tuto úlohu řešit algoritmicky. Bylo publikováno již několik detektorů HFO, ale doposud neexistuje robustní obecně fungující metoda. V rámci naší výzkumné skupiny jsou vyvíjeny dva detektory s novým přístupem. Pro měření kvalit detektorů se používají odhady statistických ukazatelů získaných porovnáním vizuálně hodnocených záznamů s výsledky automatické detekce. V prezentované studii je popsána metodika statistického hodnocení, používané referenční množiny vytvořené metodami objektivizace hodnocení většího počtu hodnotitelů a výsledky statistických ukazatelů jednotlivých implementovaných detektorů.

1 Úvod

Epilepsie je rozšířené neurologické onemocnění, kterým trpí celosvětově přibližně 1,5% populace. Toto onemocnění se může projevit spontánně kdykoliv v průběhu života, může být způsobeno druhotně jiným onemocněním mozkové tkáně nebo se může vyvinout následkem poranění. Léčba pomocí medikace umožňuje u většiny léčených pacientů kontrolovat projevy onemocnění. U jedné třetiny případů však není léčba pomocí medikamentů (kombinací alespoň dvou antiepileptik) úspěšná. V takových situacích se dále zvažuje chirurgická léčba, tedy chirurgický zásah, při kterém se odstraní část mozkové tkáně, která epileptické záchvaty způsobuje. Pokud je pacient vhodným kandidátem pro tento způsob léčby, podstoupí v rámci předoperačního vyšetření řadu klinických vyšetření. Jedná se například o zobrazovací metody jako je magnetická rezonance (MRI), výpočetní tomografie (CT), pozitronová emisní tomografie atd. V případě, že jsou nálezy těchto neinvazivních vyšetření neprůkazné (nejednoznačné), ale resekcí zákrok je stále zvažován, pacientovi se dočasně implantují nitrolební intrakraniální/subdurální elektrody a je dlouhodobě monitorován po dobu několika dnů. Záznam intrakraniálního elektroencefalografu (iEEG) je kvalitnější a není zatížen takovým množstvím artefaktů jako klasické skalповé elektroencefalogramy (EEG). Implantované elektrody umístěné v přímém kontaktu s mozkovou tkání navíc umožňují především přesnější prostorovou lokalizaci epileptogenního ložiska a epileptogenní tkáně, tedy centra, které pravděpodobně záchvaty spouští resp. tkáně, která generuje patologickou aktivitu. Mezi projevy epilepsie se v EEG záznamu řadí celé množství grafoelementů. Typickým příznakem jsou například epileptogenní výboje [1]. Tyto grafoelementy se vyskytují jak v průběhu záchvatu (iktálně), tak v době mezi záchvaty (interiktálně).

Další grafoelementy, které patří mezi projevy epilepsie, jsou patologické vysokofrekvenční oscilace (High Frequency Oscillations, HFO). Stejně jako většina jiných projevů se mohou u různých typů onemocnění objevovat jak v interiktální, tak iktální fázi. Jedná se o oscilace o frekvencích v pásmu 80 – 500 Hz [2], které jsou superponovány na běžnou mozkovou aktivitu, tzv. pozadí. Tyto oscilace mohou mít velmi rozdílnou délku trvání i amplitudu. Bylo zjištěno, že se objevují v oblastech, které jsou epilepticky aktivní a mají tak pro určení ložiska dobré lokalizační schopnosti [2-4]. Aby mohla být tato jejich vlastnost správně využita je potřeba tyto úseky v dlouhodobém iEEG záznamu identifikovat. Pro lidského hodnotitele je tento úkol časově velice náročný [5] a výsledek hodnocení záznamu je navíc silně zatížen zkušenostmi hodnotitele. Díky tomu se vizualní hodnocení stává subjektivním. Je tudíž potřeba tuto úlohu provádět nezávisle na lidských schopnostech pomocí automatických detektorů. Tyto detektory jsou založeny na signálovém zpracování a statistickém hodnocení. V současnosti již bylo publikováno několik detektorů s podobným zaměřením [4, 6-7].

Jelikož stále výsledky těchto algoritmů nejsou zcela přesvědčivé pro univerzální použití, domníváme se v našem výzkumném týmu, že je možné zvolit vhodnější přístup, který by byl více robustní. V rámci týmu jsou v současnosti vyvíjeny dva nezávislé detektory HFO založené na odlišných principech.

Aby bylo možné kvantifikovat úspěšnost detektorů je zapotřebí testovací množina. Tou zpravidla jsou hodnocení záznamu od většího počtu expertů. Jelikož je potřeba subjektivní hodnocení jednotlivců alespoň částečně objektivizovat, vytvoří se různými metodami tzv. hodnocení superhodnotitele.

2 Metodika

2.1 Hodnocené záznamy

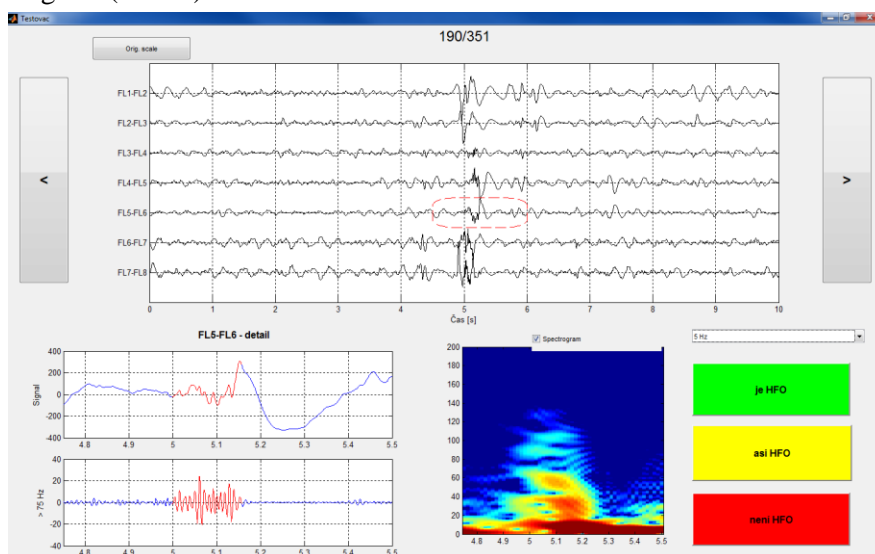
Pro předběžné porovnání detektorů byly vybrány záznamy od dvou dětských pacientů, v jejichž iEEG záznamech byly vizuálně rozpoznatelné HFO události. Tito pacienti byli v rámci předoperačního vyšetření monitorováni na klinice dětské neurologie centra pro epilepsii Motol. Elektrody byly implantovány před plánovaným resekcčním zákrokem a nahrávání započalo den po implantaci, kdy se pacienti vrací na pokoj určený k pořízení synchronizovaného EEG a audiovizuálního záznamu (video-EEG). U obou pacientů byla použita kombinace hlubkových intrakraniálních makroelektrod a povrchových subdurálních makroelektrodoých pásků (strip) s různou prostorovou lokalizací v závislosti na předpokládané lokalizaci ložisek. Záznam byl pořízen se vzorkovací frekvencí 1000 Hz v referenčním zapojení.

U pacienta č. 1 bylo nahráváno 97 svodů, u pacienta č. 2 pak 94 svodů. Pro podrobné hodnocení a označení jednotlivých HFO událostí byly na základě rutinního vizuálního hodnocení vybrány kanály, ve kterých byly HFO přítomné a kanály, které se s těmito nacházejí na stejném elektrodoém pásku. Hodnocen byl od každého pacienta jeden záznam z interiktální fáze (p. č. 1: 3,3 min., p. č. 2: 9 min.). U pacienta č. 1 bylo hodnoceno 16 elektrod (2 pásy po 8 kontaktech), u pacienta č. 2 bylo hodnoceno 8 elektrod (1 pásek po 8 kontaktech).

2.2 Vizuální hodnocení

Pro veškerá hodnocení bylo použito zobrazení signálů v bipolární montáži. Na základě vizuálního hodnocení záznamů byly laikem vybrány úseky s podezřelou vysokofrekvenční aktivitou. Byly označovány všechny úseky, u kterých nebylo zřejmé, že se jedná o běžnou aktivitu pozadí. Celkově bylo vybráno 593 událostí (242 u pacienta č. 1, 351 u pacienta č. 2).

Takto předvybrané události byly následně hodnoceny skupinou 4 hodnotitelů (2 neurologové, 2 laici). Pro hodnocení byla v prostředí Matlab vytvořena GUI aplikace. Zobrazeno bylo 10 s signálů několika kanálů s označením podezřelého úseku a možností zapnutí filtrace horní propustí (5 nebo 75 Hz). Dále se zobrazoval detail úseku se zvětšenou amplitudovou i časovou osou, detail filtrovaného signálu a spektrogram (Obr. 1).



Obr. 1: Ukázka aplikace pro hodnocení podezřelých úseků.

Pro ohodnocení daného úseku mohla být zvolena jedna z možností („Je HFO“, „Asi je HFO“, „Není HFO“). Snahou hodnotitelů bylo používat „Asi je HFO“ v co nejmenší míře, pouze v případech, kdy nedokázali s jistotou vyvrátit zbývající možnosti.

Před samotným hodnocením bylo prodiskutováno, jaké vlastnosti by měl úsek splňovat, aby byl považován za HFO, mezi všemi hodnotiteli. Tato diskuze proběhla za účelem zmenšit rozdíly v hodnocení mezi hodnotiteli a stanovit pro celou skupinu pravidla platná pro všechna následující hodnocení.

Na základě publikovaných postupů vizuálního hodnocení [3, 4, 7, 8] byla stanovena následující kritéria. Úsek musí mít v pásmu nad 75 Hz výrazný nárůst amplitudy (> 2 krát) proti okolí (pozadí), v původním nefiltrovaném signálu musejí být patrné oscilace sinusového charakteru a to alespoň 3 po sobě jdoucí periody, ve spektrogramu musí nárůst energie významně zasahovat do pásma nad 60 Hz a musí být jasně stanovitelný počátek a konec úseku.

2.3 Porovnávání detektory

Pro srovnání schopnosti automaticky detekovat úseky HFO byly vybrány kromě dvou námi vyvíjených detektorů také tři konkurenční detektory, které byly již dříve publikovány. Tyto detektory byly implementovány v jazyce Matlab na základě popisu uváděného autory. U detektorů nebyly prováděny žádné úpravy a vstupní parametry byly zvoleny v souladu s doporučením autorů

První implementovaný detektor je založen na změnách energie. Jedná se o nejstarší publikovaný detektor určený k označování HFO [4]. Algoritmus počítá efektivní hodnotu (Root Mean Square, RMS) filtrovaného signálu (100-500 Hz) v klouzavém okně. Úseky, kde je hodnota RMS větší než pětinašobek směrodatné odchylky nad průměrem z celého záznamu po dobu alespoň 6 ms, jsou označeny jako předpokládané HFO. Dodatečná pravidla kontrolují, zda absolutní hodnota filtrovaného signálu v úseku splňuje požadavek alespoň 6 vrcholů větších než trojnásobek směrodatné odchylky nad průměrnou hodnotou. Tento detektor bude dále označován jako *RMS Detector* [4, 9].

Další detektor používá krátkodobé délky čáry (short-time line length) pro identifikaci úseků podezřelých z přítomnosti HFO. Nejprve je signál filtrován do pásma 80-100 Hz a následně je spektrálně ekvalizován diferenčním filtrem prvního řádu. Práhová hodnota je odhadnuta jako percentil 97,5 kumulativní distribuční funkce pro 3 minuty dlouhé úseky. V algoritmu nejsou používány žádné další kroky pro korekci. Dále bude tento detektor označován jako *LineLength Detector* [6].

Poslední z implementovaných konkurenčních detektorů používá výpočet Hilbertovi obálky z filtrovaného signálu (180-400 Hz). Obálka je prahována hodnotou pětinašobku směrodatné odchylky vypočtené z celého záznamu. Za identifikované události jsou označeny vždy nadpráhová lokální maxima obálky. Dále bude detektor označován jako *Hilbert Detector* [7].

První z detektorů navržený a implementovaný v naší výzkumné skupině také používá výpočet Hilbertovy obálky filtrovaného signálu (80-500 Hz). Hodnota, kterou je obálka prahována se dynamicky upravuje na základě odhadu amplitudové distribuce v jednotlivých segmentech délky 4 sekundy. Hodnota prahu je váhovanou kombinací mediánu a střední hodnoty této distribuce. Dodatečně jsou pak spojeny úseky, které jsou vzájemně bližší než 40 ms a odstraněny úseky kratší než 50 ms. Detektor bude nadále označován jako *ISARG H. Detector* [10].

Druhý navržený a implementovaný detektor využívá pro prvotní segmentaci změnu krátkodobé energie (okno 100 ms) v pásmu nad 60 Hz. Jako zajímavé segmenty jsou brány ty, které přesáhnou percentil 95 stanovený z úseku dlouhého 90 s. Dále jsou u detekovaných úseků sledovány další frekvenční i časové parametry. Ve frekvenční oblasti je požadován významný nárůst energie v pásmu nad 75 Hz ve Welchově spektru. Z časové oblasti je získán počet lokálních extrémů s významnou změnou gradientu v jejich okolí. Pokud je v daném segmentu požadovaný nárůst energie a alespoň 9 lokálních extrémů, předpokládá se, že se jedná o úsek s HFO. Detektor bude nadále označován jako *ISARG B. Detector*.

2.4 Statistické hodnocení

Z hodnocení jednotlivých hodnotitelů byly za účelem objektivizace vytvořeny množiny hodnocení tzv. superhodnotitelů. Byly zvoleny tři různé metody pro vytvoření těchto souborů. Díky tomu je dosaženo větší výtěžnosti informace, což je vzhledem k relativně malému vzorku dat výhodné.

Pro potřeby algoritmického zpracování jsou pro jednotlivé možnosti hodnocení použity následující číselné ekvivalenty: „Je HFO“ = 1, „Asi je HFO“ = 0,5, „Není HFO“ = 0.

První z použitých souborů hodnocení pro ověření vlastností detektorů je testovací množina nazvaná *Ground Truth* (GT). V této množině jsou za úseky HFO uznány pouze ty, které všichni hodnotitelé označili „Je HFO“ [6]. Matematicky popsaná podmínka úseku splňující požadavky na HFO v množině GT je následující:

$$\sum_{i=1}^N x_i = N \quad (1)$$

kde x_i jsou jednotlivá hodnocení jednoho úseku, N je celkový počet hodnotitelů. Jedná se tedy o nejprísnější podmínku, pro stanovení úseku HFO. Další běžný postup vytvoření množiny superhodnotitele je *Gold Standard* (GS) [11]. Za HFO jsou označeny úseky, na kterých se shodla více jak polovina hodnotitelů. Matematicky zapsaná podmínka:

$$\sum_{i=1}^N \frac{x_i}{N} > \frac{N}{2} \quad (2)$$

Poslední z použitých množin se nazývá *Ground Possibly* (GP). V této množině jsou nejméně přísná kritéria pro označení úseku jako HFO. Reflektuje úseky, o kterých se alespoň jeden z hodnotitelů domnívá, že se jedná o HFO nebo alespoň dva hodnotitelé si nejsou jistí, že se s jistotou o HFO nejedná. Matematicky je vztah pro určení úseku jako HFO následující:

$$\sum_{i=1}^N x_i \geq 1 \quad (3)$$

Množina GP slouží především pro kontrolu, zda detektory identifikují i rozporuplné úseky nebo jsou zaměřeny především na detekci jasnějších případů. Hodnocení detektorů s použitím této množiny je spíše doplňující informací.

Pro statistické hodnocení schopnosti detektorů identifikovat v záznamu HFO je potřeba ošetřit situace, kdy se detekovaný úsek nepřekrývá přesně s úsekem, který byl vizuálně označen. K takovým situacím dochází velmi často. Vizualně hodnocené úseky byly rozšířeny o 10 ms proti i po směru běhu času a jejich správné označení bylo uznáno, pokud se s detekovaným úsekem překrývaly alespoň jedním vzorkem (viz. Obr. 2). Za skutečně pozitivní detekci (*True Positive*, TP) byl považován vždy jen jeden z překrývajících se detekovaných úseků. V případě, že se s jedním vizuálně označeným úsekem HFO překrývala více jak jedna detekce, ostatní byly považovány za falešně pozitivní detekce (*False Positive*, FP).



Obr. 2: Ilustrace hodnocení detekcí jako TP/FP. Zelené úseky – TP, červené úseky – FP, černá tlustá čára – vizuálně označený úsek HFO, tečkovaně – rozšíření označení HFO

Pro porovnání detektorů byly použity dva statistické ukazatele. Senzitivita (*Sensitivity*, Sen):

$$Sen = \frac{TP}{TP + FN} [\%] \quad (4)$$

kde TP vyjadřuje celkový počet skutečně pozitivních detekcí, FN je celkový počet falešně negativních detekcí. Ukazatel vyjadřuje procento detekovaných HFO ze všech úseků označených jako HFO v testovací množině. Pozitivní prediktivní hodnota (*Positive Predictive Value*, PPV):

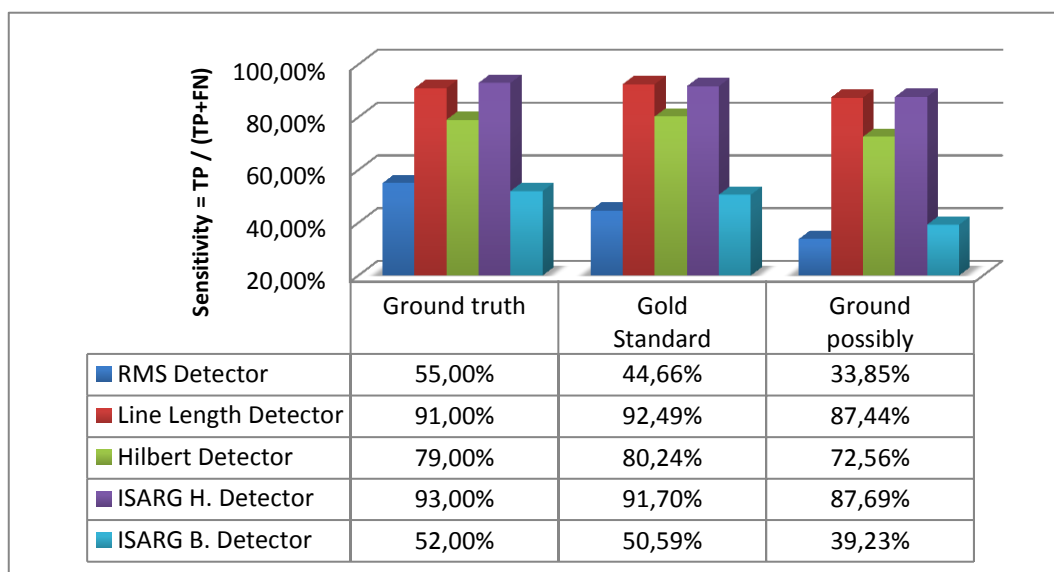
$$PPV = \frac{TP}{TP + FP} [\%] \quad (5)$$

kde *FP* vyjadřuje celkový počet falešně pozitivních detekcí. Tento ukazatel vyjadřuje, poměr kolik z detekovaných úseků bylo zároveň v testovací množině označeno jako HFO. Další běžně používaný statistický ukazatel specificita nemůže být použit, neboť u tohoto typu dat není možné stanovit počet skutečně negativních (*False Negative*, FN) úseků.

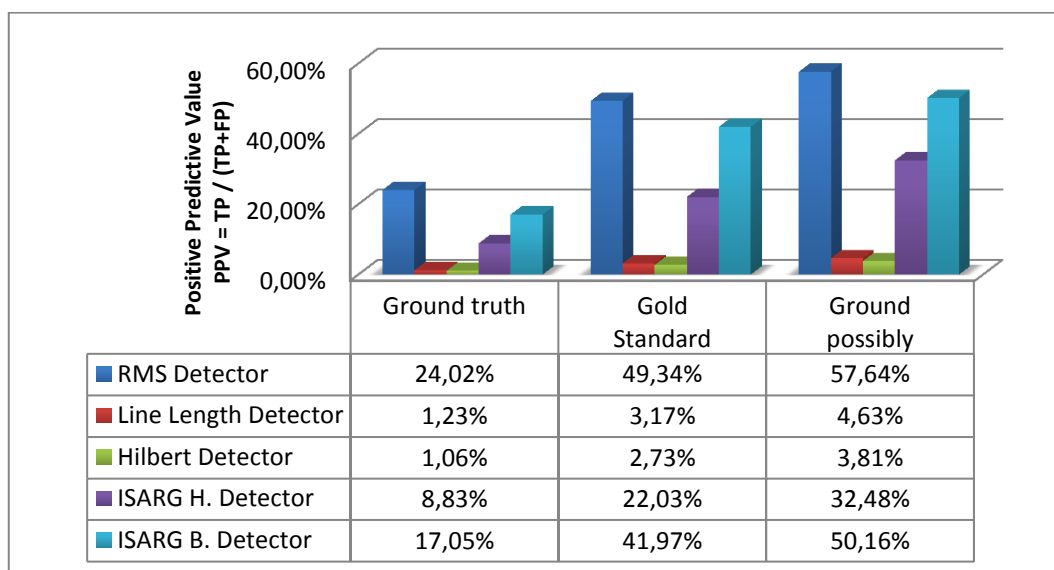
3 Výsledky

Použité množiny superhodnotitele mají rozdílné počty událostí považovaných za HFO v závislosti na zvolených kritériích. U množiny Ground Truth (GT) bylo pouze 100 úseků uznáno jako HFO, tedy ohodnoceno „Je HFO“ všemi hodnotiteli. V testovací množině Gold Standard (GS) bylo jako HFO uznáno 253 úseků a v množině Ground Possibly (GP) pak 390 úseků.

Po algoritmickém zpracování signálů všemi implementovanými detektory byly zaznamenány také velmi odlišné počty detekovaných úseků jednotlivými detektory. RMS Detector celkově provedl 229 detekcí, LineLength Detector 7371, Hilbert Detector 7432, ISARG H. Detector 1053 a ISARG B. Detector 305. Číselné hodnoty dosažené senzitivity a grafické ilustrace jsou uvedeny na Obr. 3 a pozitivní prediktivní hodnoty jsou uvedeny na Obr. 4.



Obr. 3: Senzitivita dosažená jednotlivými detektory při zvolení rozdílných testovacích množin



Obr. 4: Pozitivní prediktivní hodnota dosažená jednotlivými detektory při zvolení rozdílných testovacích množin

4 Závěr a diskuze

Pro účely statistického hodnocení vlastností detektorů je vhodné, aby testovací data byla co nejrozsáhlejší a obsahovala vzorky od většího počtu subjektů, aby odhady statistických ukazatelů nebyly významně vychýleny. Primárním cílem tohoto předběžného hodnocení kvality detektorů byla konfrontace výsledků námi navržených detektorů se známými publikovanými detektory na vlastním dostupném vzorku dat se známými vlastnostmi. Doposud dostupná data nejsou požadovaného rozsahu, aby bylo možné považovat výsledky za významné. Délka hodnocených záznamů, přepočteno na jeden kanál, činí přibližně 54 minut (pacient č. 1) a 72 minut (pacient č. 2). Prezentované výsledky jsou orientační a slouží autorům detektorů při jejich dalším vývoji.

Vizuální hodnocení, kdy hodnotitel hledá v záznamu úseky s výskytem HFO, je časově extrémně náročné (hodnocení 10 minut, 10 kanálů trvá až 10 hodin [5]). Proto byl zvolen postup, při kterém byly předvybrány podezřelé úseky, které následně ohodnotil větší počet hodnotitelů. Tento postup připouští chybu kromě vlastního hodnocení i ve fázi předvýběru. Z toho důvodu jsou za účelem minimalizace této chyby předvybrány i úseky s minimálním podezřením na přítomnost HFO. Hodnotitelé jsou tímto způsobem schopni ohodnotit průměrně 45 předvybraných událostí za 10 minut.

Metody vytváření testovacích souborů superhodnotitele (metody objektivizace) ze subjektivních hodnocení většího počtu hodnotitelů významně ovlivňují odhady statistických ukazatelů. Pro test na stejném souboru záznamů poskytují jednotlivé množiny (GT, GS, GP) různou referenci, se kterou jsou srovnávány výsledky algoritmických detekcí. Pravidla, na základě kterých se množiny vytvářejí, různě upravují schopnosti superhodnotitele. Statisticky vzato balancují senzitivitu a selektivitu s jakou superhodnotitel HFO identifikuje. Množina Ground Possibly byla v této studii použita s cílem zjistit, zda jsou detektory citlivé na sporné úseky nebo jejich nadbytečné detekce oproti množinám GT a GS detekují nerelevantní události, což by se projevilo výraznějším poklesem senzitivity.

Při srovnávání detektorů je patrné, že LineLength Detector a Hilbert Detector mají několikanásobně větší počet detekcí než ostatní detektory. Tento jev měl kladný dopad na odhadovanou senzitivitu, ale při hodnocení PPV pochopitelně nedosahují nejlepších hodnot. Autoři uvádějí, že je cílem publikovaných detektorů dosahovat co nejlepší senzitivity i za cenu zvýšeného počtu falešných detekcí. Při příliš zvýšeném počtu falešných detekcí přestávají však být výsledky identifikace použitelné pro další zpracování, např. pro kvantitativní hodnocení a lokalizaci epileptogenní zóny.

Konkurenční detektory byly implementovány na základě publikovaného popisu a nebyly u nich prováděny dodatečné úpravy nebo optimalizace. Jelikož byly detektory optimalizovány na rozdílných datových souborech nahraných a zpracovávaných jinými EEG systémy je možné, že by po optimalizaci parametrů pro nám dostupné typy záznamů došlo ke zlepšení jejich kvalit. Obdobné srovnání některých detektorů ve svých pracích uvádějí také Zelman a spol. [12] a Gardner a spol. [6], kteří taktéž konfrontovali publikované detektory s vlastními návrhy. Rozdíly v použitých databázích signálů ovlivňují schopnosti jednotlivých detektorů v takových záznamech detekovat HFO.

Poděkování

Tato práce byla podporována Ministerstvem zdravotnictví České Republiky prostřednictvím grantů IGA NT11460-4/2010, IGA NT13357-4/2012 a Studentskou grantovou soutěží ČVUT v Praze SGS 10/272/OHK4/3T/13.

Reference

- [1] STEAD, Matt, et al. Microseizures and the spatiotemporal scales of human partial epilepsy. *Brain*, 2010, 133.9: 2789-2797.
- [2] BRAGIN, Anatol, et al. Interictal high-frequency oscillations (80–500Hz) in the human epileptic brain: Entorhinal cortex. *Annals of neurology*, 2002, 52.4: 407-415.
- [3] BRAGIN, Anatol, et al. Hippocampal and entorhinal cortex high-frequency oscillations (100–500 Hz) in human epileptic brain and in kainic acid-treated rats with chronic seizures. *Epilepsia*, 1999, 40.2: 127-137.

- [4] STABA, Richard J., et al. Quantitative analysis of high-frequency oscillations (80–500 Hz) recorded in human epileptic hippocampus and entorhinal cortex. *Journal of neurophysiology*, 2002, 88.4: 1743-1752.
 - [5] ZELMANN, R., et al. Automatic detector of high frequency oscillations for human recordings with macroelectrodes. In: *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. IEEE, 2010. p. 2329-2333.
 - [6] GARDNER, Andrew B., et al. Human and automated detection of high-frequency oscillations in clinical intracranial EEG recordings. *Clinical neurophysiology*, 2007, 118.5: 1134-1143.
 - [7] CRÉPON, Benoît, et al. Mapping interictal oscillations greater than 200 Hz recorded with intracranial macroelectrodes in human epilepsy. *Brain*, 2010, 133.1: 33-45.
 - [8] SALAMI, Pariya, et al. A COMPARISON BETWEEN AUTOMATED DETECTION METHODS OF HIGH-FREQUENCY OSCILLATIONS (80-500Hz) DURING SEIZURES. *Journal of Neuroscience Methods*, 2012.
 - [9] STABA, Richard J., et al. High-frequency oscillations recorded in human medial temporal lobe during sleep. *Annals of neurology*, 2004, 56.1: 108-115.
 - [10] HAVEL, T., et al. Automatic detection of high-frequency oscillations in invasive recordings. In: *Medical Measurements and Applications Proceedings (MeMeA), 2013 IEEE International Symposium on*. IEEE, 2013. p. 228-232.
 - [11] WEBBER, W. R. S., et al. Practical detection of epileptiform discharges (EDs) in the EEG using an artificial neural network: a comparison of raw and parameterized EEG data. *Electroencephalography and clinical Neurophysiology*, 1994, 91.3: 194-204.
-

Ing. Tomáš Havel
Email: havelto3@fel.cvut.cz

Ing. Petr Ježdík, Ph.D.
Email: jezdipl@fel.cvut.cz