

Užití statistických modelů pro ztrátovou agregaci dat kvality elektrické energie

Tomáš Bedrník <tomas.bedrnik@tul.cz>, Leoš Kukačka leos.kukacka@tul.cz
<tomas.bedrnik@tul.cz>, Ekaterina Nyrobotseva <ekaterina.nyrobotseva@tul.cz>, Vratislav Žabka
<vratislav.zabka@tul.cz>

Příspěvek popisuje návrh algoritmu na ztrátovou agregaci specifických dat pocházejících z široké kampaně na sběr dat týkajících se kvality elektrické energie. Algoritmus využívá dynamického dělení dat na časové úseky, ve kterých lze data s výhodou popsat běžnými známými rozděleními hustoty pravděpodobnosti.

Klíčová slova: power quality, ztrátová komprese, agregace, normální rozdělení

Úvod

Analyzátoři kvality energie nebo chytré elektroměry jsou stále častěji nasazovány v průmyslu i domácnostech. Množství dat, které tyto přístroje ukládají, proto velice rychle roste.

Práce se zabývá ztrátovým algoritmem pro agregaci dat z databáze měření kvality elektrické energie (PQ). Zdrojová data představují RMS hodnoty napětí získávané po půl periodě základního síťového kmitočtu. Cílem je vyvinout algoritmus, který nabízí vyšší kompresní poměr při zachování nebo vylepšení informační hodnoty komprimovaných dat oproti standardnímu přístupu [1].

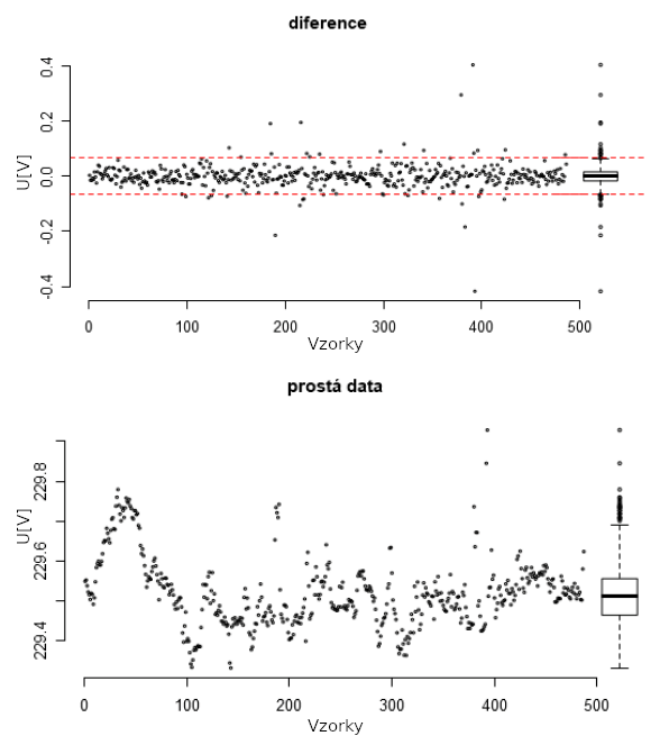
Metodika

Nějakou formu bezztrátové komprese dat používá většina analyzátorů a elektroměrů. Možnosti zlepšování v těchto algoritmech jsou už téměř vyčerpány, proto je nutné přistoupit ke ztrátové kompresi nebo agregaci dat. Existuje mnoho přístupů - waveletové transformace [2], fuzzy logika [3], automatické učení, selektivní agregace na základě indexu kvality elektrické energie [4] a mnoho dalších.

V této práci je prezentován nový, ještě nepopsaný přístup - použití statistických metod k redukci ukládaných dat. Místo aby se ukládaly hodnoty sledované veličiny v pravidelných intervalech, tak se použije statistické rozdělení pravděpodobnosti k popisu delších úseků dat, které mu vyhovují.

Výsledky a diskuze

Algoritmus pracuje s dynamickým rozdělením intervalů a oproti standardnímu přístupu, kdy se ze zdrojových dat ukládá pouze průměrná hodnota, minimum a maximum, testuje hypotézu, zda lze v daném intervalu zdrojová data modelovat Gaussovým rozdělením. V takovém případě zaznamená rovněž směrodatnou odchylku.



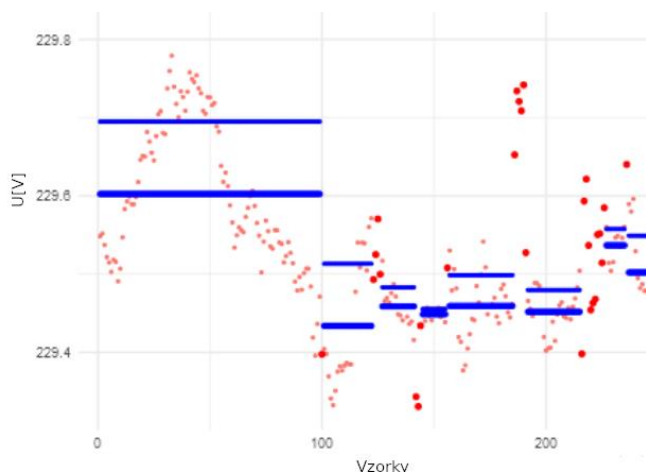
Obrázek 1: Zdrojová data a první diference

V prvním kroku se vypočítá první diference hodnot časové řady a vybere odlehlé hodnoty (znázorněno na obr. 2 červenou čárkovanou čarou). Pomocí nalezených

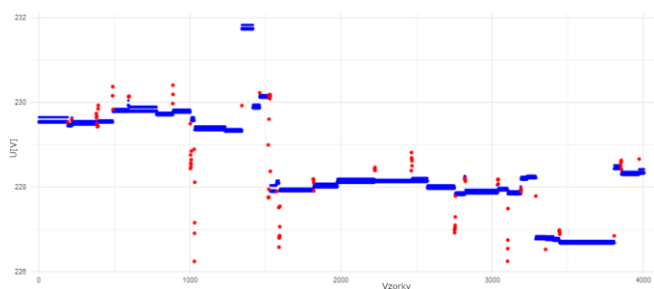
odlehých hodnot rozdělí řadu na intervaly, přičemž odlehle hodnoty nezahrnuje do intervalů. Dále spočítá, zda difference či data v jednotlivých intervalech splňují kritéria pro vybrané rozdělení.

V druhém kroku algoritmus vybere pouze ty intervaly, které nesplňují kritéria daného rozdělení, ale které jsou zároveň delší než zadaný počet prvků (default 8). S použitím algoritmu jedna dále rozděluje úseky tak, aby výsledný soubor dat obsahoval pouze intervaly splňující kritéria rozdělení nebo samostatné body.

Ve třetím kroku má algoritmus za úkol minimalizovat počet intervalů a počet samostatných bodů. Postupně testuje spojování intervalů se sousedními samostatnými body či intervaly takovým způsobem, aby výsledný soubor dat obsahoval stále pouze intervaly splňující kritéria zvoleného rozdělení. Také nesmí sloučením sousedních intervalů dojít k významné ztrátě přesnosti.



Obrázek 2: Výsledek – sytě červeně samostatné body, světle původní data a modře úseky normálního rozdělení



Obrázek 3: Výsledek na delším úseku dat

První tři grafy odpovídají krátkému úseku dat a ilustrují rozdělování řady na intervaly a samostatné body. Pro intervaly jsou vykresleny dvě modré čáry odpovídající průměru a k němu přičtené směrodatné odchylce. Velké červené body značí samostatné hodnoty vhodné k uložení, malé červené body pak původní data.

Poslední graf je výsledkem popsanych tří algoritmů pro delší úsek dat o 4000 prvcích. V tomto případě bylo vytvořeno 40 intervalů a 118 samostatných bodů.

Závěr

Práce popisuje algoritmus pro ztrátovou agregaci dat analyzátorů PQ, který pracuje na principu nahrazování souvislých úseků dat pomocí Gaussova rozdělení.

Algoritmus byl úspěšně implementován a otestován, takže je ověřena jeho funkčnost. Práce je ale teprve na počátku. Je nutné ověřit přínosy algoritmu - kompresní poměr a přesnost popisu dat.

Poděkování

Tato práce byla podpořena z projektu Studentské grantové soutěže (SGS) na Technické univerzitě v Liberci v roce 2019.

Reference

- [1] IEC 61000-4-30: Electromagnetic compatibility (EMC) - Part 4-30: Testing and measurement techniques - Power quality measurement methods. 2015.
- [2] BRINDHA, S. a D. SUNDARARAJAN, 2013. Power quality monitoring and compression using the discrete wavelet transform. In: 2013 International Conference on Advanced Computing and Communication Systems: 2013 International Conference on Advanced Computing and Communication Systems [online]. s. 1–6.
- [3] IBRAHIM, W. R. A. a M. M. MORCOS, 2005. Novel data compression technique for power waveforms using adaptive fuzzy logic. IEEE Transactions on Power Delivery [online]. 20(3), 2136–2143. ISSN 0885-8977.
- [4] Bedrník, T.; Kukačka, L.; Štěpán, P.; aj.: Dynamic Intelligent Compression for Power Quality Analysers. In The 23rd International Conference and Exhibition on Electricity Distribution, ročník 23rd, Lyon, France, červen 2015.