Posudek oponenta bakalářské práce

předložené na Matematicko-fyzikální fakultě Univerzity Karlovy

Autor:	Petr Sedláček
Název práce:	Limitations of incompressible encodings
Stud. program a obor:	matematika, obecná matematika
Rok odevzdání:	2021
Jméno a tituly oponenta:	Mgr. Martin Mareš, Ph.D.
Pracoviště:	Katedra aplikované matematiky
Kontaktní e-mail:	mares@kam.mff.cuni.cz

Shrnutí obsahu práce

Předložená práce se zabývá nekomprimovatelným kódováním dat. To je randomizované kódování s následujícími vlastnostmi: (1) zakódované zprávy nejsou příliš dlouhé (co to přesně znamená, je dáno nějakým parametrem), (2) zakódované zprávy lze s vysokou pravděpodobností bezztrátově dekódovat, (3) ke znovuvytvoření zakódované zprávy je kromě originální zprávy potřeba netriviální množství dat (určeno dalším parametrem).

Moran a Wichs v roce 2020 popsali konstrukci takového kódování založenou na hypotézách z teorie složitosti, která splňuje vlastnost (3) pro polynomiálně omezeného protivníka. Ve svém článku zmiňují tvrzení, že nepodmíněně korektní kódování (které by připouštělo neomezeně silného protivníka) nemůže existovat. Důkaz je ale pouze naznačen a má četné slabiny.

Cílem této práce bylo tvrzení dokázat rigorozně. Tohoto cíle student dosáhl. V první části práce definuje nekomprimovatelné kódování a zavádí značení. Poté vysvětluje problémy původního důkazu.

Nakonec uvádí vlastní důkaz, který je vystavěn postupně: Nejprve uvažuje případ, kdy je dekódování deterministické zcela bezztrátové – tehdy stačí jednoduchý počítací argument. Pak jednoduchou úpravou povoluje ztrátovost dekódování. V posledním kroku se vyrovnává s možným nedeterminismem dekódovacího algoritmu, což už vyžaduje složitější úvahy.

Celkové hodnocení práce

- *Téma* patří mezi náročnější a vyžaduje pochopení netriviálních konceptů z kryptografie nad rámec běžného bakalářského studia. Práce jistě splňuje zadání.
- *Vlastním příspěvkem* autora je důkaz tvrzení, které bylo již dříve uvedeno v literatuře, ale dokázáno chybně.
- *Matematická úroveň* práce je výborná. Tvrzení jsou exaktně formulována a dokázána.
- *Práce se zdroji.* Použité zdroje jsou správně citovány.

- *Formální úprava.* Práce je psaná čtivou angličtinou bez jazykových a typografických chyb. Občas se vyskytují překlepy v matematické notaci (viz příloha). Tím spíš by srozumitelnosti práce pomohlo, kdyby před formální definicí nebo tvrzením byl alespoň naznačen neformální pohled.

Závěr

Práci považuji za kvalitní a doporučuji ji uznat jako bakalářskou.

V Praze dne 21. června 2021
Martin Mareš

Příloha: drobné chyby

- Definice 2.1: Bylo by vhodné uvést, co je vstupem a výstupem algoritmů.
- Definice 2.4: Bylo by dobré říci, jakého typu jsou α a β (tedy že to nejsou konstanty, nýbrž funkce, a co znamenají jejich argumenty). V definici experimentu též specifikovat, zda algoritmy Select, Compress a Expand jsou deterministické nebo randomizované.
- Textu by velice prospělo uvést význam parametrů α a β i neformálně.
- Strana 4, definice D_m: zde je trochu zavádějící psát rovnost $\mathrm{Dec}(c) = m$, když algoritmus Dec je randomizovaný.
- Strana 7, definice $\widehat{\mathrm{Enc}}$: zde je potřeba, aby funkce q byla efektivně vyčíslitelná (v pravděpodobnostním polynomiálním čase). V daném kontextu není problém takovou funkci sehnat, ale je na to potřeba dávat pozor.
- Lemma 4.10: V tvrzení i důkazu je popleteno m a m^*.
- Lemma 4.11: Někde se indexuje od 1 do n, jinde tentýž objekt od 0 do n. Tvrzení předpokládá klesající posloupnost $\mathbf{y}$, ale celá druhá polovina důkazu se zabývá případem, kdy posloupnost je nerostoucí. Další použití lemmatu přitom zaručuje pouze nerostoucí $\mathbf{y}$. Předpokládám tedy, že je to překlep v tvrzení. Pak už ale není nutně splněno, že $\mathbf{z}$ je jediný vektor, kde se maxima skalárního součinu $\langle \mathbf{x}, \mathbf{y} \rangle$ nabývá. Korektnosti dalších úvah to nicméně opět nebrání.
- Strana 13, řádek začíající „Next, we sort“: místo „$\cdots$“ má být „$\ldots$“.
- Strana 14, 4. řádek zdola: „if“ nemá být sázeno matematickou kurzívou.
- Typografie: Některé jednopísmenné proměnné (zejména p a c) občas nejsou sázeny matematickou kurzívou.