

Slabiková komprese je nový přístup ke kompresi textů po symbolech. Důležitým aspektem tohoto přístupu jsou slovníky častých slabik. Tyto slovníky slouží k počáteční inicializaci kompresních algoritmů a mají velký vliv na kompresní poměr. Doposud byly získávány přímočaře na základě četností výskytů slabik v korpusech. Věříme, že za pomoci genetického algoritmu bychom mohli dokázat přesněji určit množinu slabik, které jsou pro daný jazyk charakteristické, a tím dosáhnout lepších výsledků při samotné kompresi. V této práci se pokusíme takovýto algoritmus navrhnout a otestujeme jej na textech v anglickém a českém jazyce.