

Word alignment (párování slov) je klíčovou komponentou moderních systémů statistického strojového překladu. Vstupem je věta ve dvou jazycích a úkolem automaticky spárovat slova v obou jazycích tak, aby se našly nejpravděpodobnější překladové ekvivalenty. Jako varianta klasického generativního přístupu (IBM modely) se dnes prosazují i diskriminativní přístupy, které tuto úlohu řeší jako hledání maximálního hranového pokrytí v úplném ohodnoceném bipartitním grafu. Vrcholy grafu jsou tvořeny slovy v jednom a v druhém jazyce, hrany jsou ohodnoceny mírou asociace odhadnutou z trénovacích dat. Práce se zaměřuje na efektivní implementaci algoritmu pro hledání maximálního pokrytí bipartitního grafu, implementaci výpočtu ohodnocení hran bipartitního grafu a provedení základních experimentů.