XML dokumenty a související technologie reprezentují jednu z nejrozšířenějších cest údržby a výměny dat na Webu. Velké množství reálných dokumentů ale bohužel obsahuje nejrůznější formy nekonzistence, které brání jejich úspěšnému a automatizovanému zpracování.
V této práci se konkrétně věnujeme problému strukturální nevalidity a její korekce. Máme-li tedy jeden potenciálně nevalidní XML dokument modelovaný jako strom a současně jeho schéma v jazycích DTD nebo XML Schema modelované jako regulární stromová gramatika, naším cílem je najít všechny minimální opravy tohoto stromu.
Námi navržený model využívá rekurzivně vnořovaných struktur korekčních multigrafů, ve kterých hledáme nejkratší cesty. Za tímto účelem formálně představíme tři korekční strategie s rozdílnými úrovněmi aplikovaných optimalizací. S ohledem na provedené experimenty pak konkrétně Refinement strategie nejenom významně překonává všechny ostatní existující přístupy, ale zároveň garantuje důležité charakteristiky, které jiné přístupy zaručit nemohou.