

Although XML is a widely used technology, the majority of real-world XML documents does not conform to any particular schema. To fill the gap, the research area of automatic schema inference from XML documents has emerged. This work refines and extends recent approaches to the automatic schema inference mainly by exploiting an obsolete schema in the inference process, designing new MDL measures and heuristic excluding of excentric data inputs. The work delivers a ready-to-use and easy-to-extend implementation integrated into the jInfer framework (developed as a software project). Experimental results are a part of the work.