

AFFIRMATIVE SAMPLING

MARTINEZ CONRADO

Random sampling from a data stream is a fundamental task at the core of many other procedures of data stream analysis. Affirmative sampling is a simple and elegant algorithm which delivers a random sample of S distinct elements from the stream, each of the n (unknown) distinct elements in the stream having equal probability of being sampled and—here's the main novelty—the expected size of S will grow with n , hence better capturing the "diversity" in the stream and allowing for more accurate