The External Path Length in Tries under the Markov Model

Kevin Leckey
Institute for Mathematics
J.W. Goethe University Frankfurt

joint work with Ralph Neininger

Tries are a widely used data structure on words. For a given finite set of words (of infinite length) on the binary alphabet $\{0, 1\}$ a trie is a binary tree which contains all words in its leaves and is built by the following recursive manner: If there is only one word to store the trie is a single node containing this word; if there are at least two words to store the trie consists of a root node and two subtrees in which the left subtree is constructed by picking all words that start with 0, deleting their first bit and building a trie; the right subtree is built similary by taking all words that start with 1.
Most of the analysis of tries is done under the simplest Bernoulli source model where each word is a sequence of iid coin tosses with success probability $p \in (0, 1)$. In this talk we focus on the analysis when each word is generated by a Markov chain. We use the contraction method to obtain a limit theorem for the external path length of a trie under this model.