

MSc thesis defense presentation

Γεωργία Αβαρικιτή defends her

MSc thesis.

Date:	Τετάρτη, 21 Ιούν 2017
Ώρα:	15:15
Location:	Εθνικόν και Καποδιστριακόν Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών, A56
Thesis title:	Geometric proximity problems in high dimensions
Committee:	<ul style="list-style-type: none">• Ιωάννης Εμμέρης• Αριστείδης Παγουρτζής• Βασίλειος Ζησιμπίδης

Thesis abstract

Geometric proximity problems is a class of problems in computational geometry that involve estimation of distances between geometric objects. In this work, we focus on two specific problems of this class, the computation of r -nets and the near neighbor decision problem on high dimensional spaces under the Euclidean distance, both of which are powerful tools in computational and metric geometry. Specifically, we present a new randomized algorithm which efficiently computes high dimensional approximate r -nets with respect to Euclidean distance. For any fixed $\epsilon > 0$, the approximation factor is $1 + \epsilon$ and the complexity is polynomial in the dimension and subquadratic in the number of points; the algorithm succeeds with high probability. We improve upon the best previously known (LSH-based) construction of Eppstein et al. in terms of complexity, by reducing the dependence on ϵ , provided that ϵ is sufficiently small. Moreover, our method does not require LSH but follows Valiant's approach in designing a sequence of reductions of our problem to other problems in different spaces, under Euclidean distance or inner product, for which r -nets are computed efficiently and the error can be controlled. Our result immediately implies efficient solutions to a number of geometric problems in high dimension, such as finding the $(1 + \epsilon)$ -approximate k -th nearest neighbor distance in time subquadratic in the size of the input. Additionally, we propose a new and simple data structure for the c -approximate near neighbor decision problem in high-dimensional spaces using linear space and sublinear query time for any $c > 1$: given an LSH family of functions for some metric space, we randomly project points to vertices of the Hamming cube in dimension $\leq \log n$, where n is the number of input points. The projected space contains strings which serve as keys for buckets containing the input points. The query algorithm simply projects the query point, then examines points which are assigned to the same or nearby vertices on the Hamming cube. We analyze in detail the query time for some standard LSH families.

