Boris Mirkin et al., J Comput Eng Inf Technol 2018, Volume 7
DOI: 10.4172/2324-9307-C1-020

Ordinally equivalent patterns: an approach to improving visualization in parallel coordinates

Boris Mirkin and **Alexey Myachin**
National Research University Higher School of Economics, Moscow, RF

Parallel coordinates is a well-known framework for data visualization. Rather than conventionally considering data objects as points in a multidimensional space, this approach visualizes the objects as polylines between parallel axes representing different features [1, 2]. Attempts at applying parallel coordinates to big data tables face issues. Of those probably the most important is the so-called clutter at which no data structure can be reliably seen on the picture. There have been various attempts at overcoming the issue such as edge bundling or continuous parallel coordinates. We propose a different approach: extracting ordinally-equivalent, or co-monotone, patterns. Given an ordering p =(x1, x2,…, xn) of data features, we refer to objects i1 and i2 as p–equivalent if the parallel coordinate polylines for these two objects are co-monotone, so that, for any k, k=1, 2, …, n-1: (a) xk(i1)>xk+1(i1) iff xk(i2)>xk+1(i2), (b) xk(i1)<xk+1(i1) iff xk(i2) < xk+1(i2), and (c) xk(i1) = xk+1(i1) iff xk(i2) = xk+1(i2). We refer to the two objects as ordinally equivalent if they are p-equivalent at any order p of parallel coordinates. Obviously, the ordinal equivalence is an equivalence relation, thus, correspo-nding to a partition of the observations in classes of ordinally equivalent objects. Of course, no clutter may emerge at the parallel coordinate graphs of a class of the ordinal equivalence. Although, on the first glance, finding ordinal equivalence classes is a complex, non-polynomial, combinatorial problem, we prove that, in fact, it is quadratic over n, and give several algorithms for finding the classes, depending on user's preferences. The algorithms are scalable to big data sizes. We give examples of using ordinally equivalent patterns for extraction and visualization of patterns prevailing in data related to various application domains. References Johansson, J., and Forsell, C. (2016). Evaluation of parallel coordinates: Overview, categorization and guidelines for future research. IEEE Transactions on Visualization and Computer Graphics, 22(1), 579-588. Heinrich, J., and Weiskopf, D. (2013). State of the art of parallel coordinates. In Eurographics (STARs), 95-116.

Biography

Mirkin is a Professor in the HSE RF and Professor Emeritus in the University of London UK.He has published a hundred of refereed papers and a dozen books on data analysis and decision making, made a number of key-note presentations at International conferences.

bmirkin@hse.ru; mirkin@dcs.bbk.ac.uk

Notes: