ScagnosticsJS: Scatterplot Visual Features for the Web Supplementary Document

Submission ID: short1044

1. Scagnostics measures and the exemplar plots they target

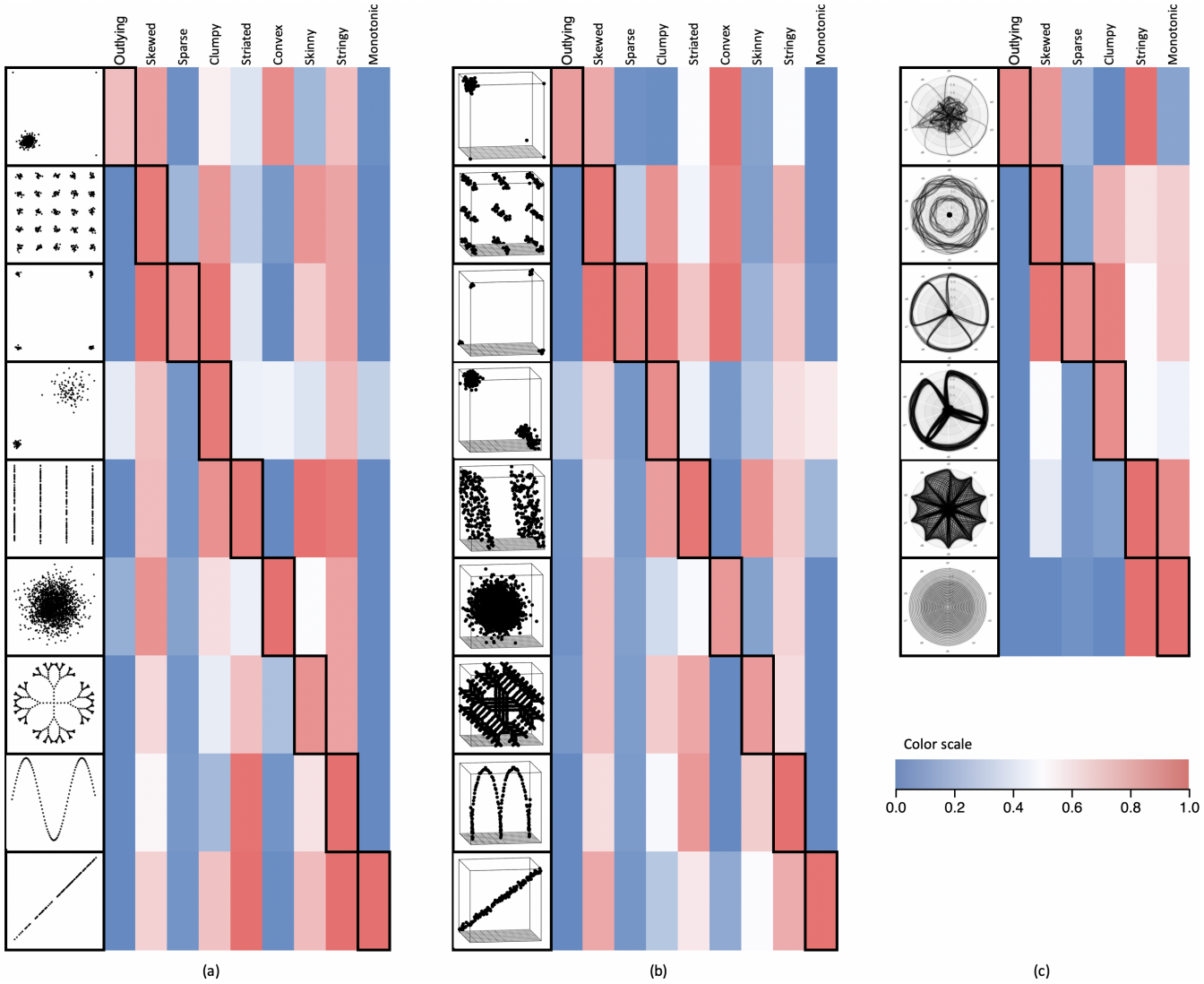

Figure 1: Scagnostics measures and exemplar plots that they target for 2D (a), 3D (b), and nD (c).

Figure 1 shows the results of our Scagnostics measures applied to the patterns of scatterplots that they target for two-dimension (2D), three-dimension (3D), and higher dimension (nD) Scagnostics JavaScript implementations. The color scale is a linear one ranging from blue (for the minimum score of 0.0) to red (for the maximum score of 1.0). It is observable that the measurements are all close to red (i.e., their corresponding values are closer to 1.0) for the patterns that they target (the diagonal cells with borders).

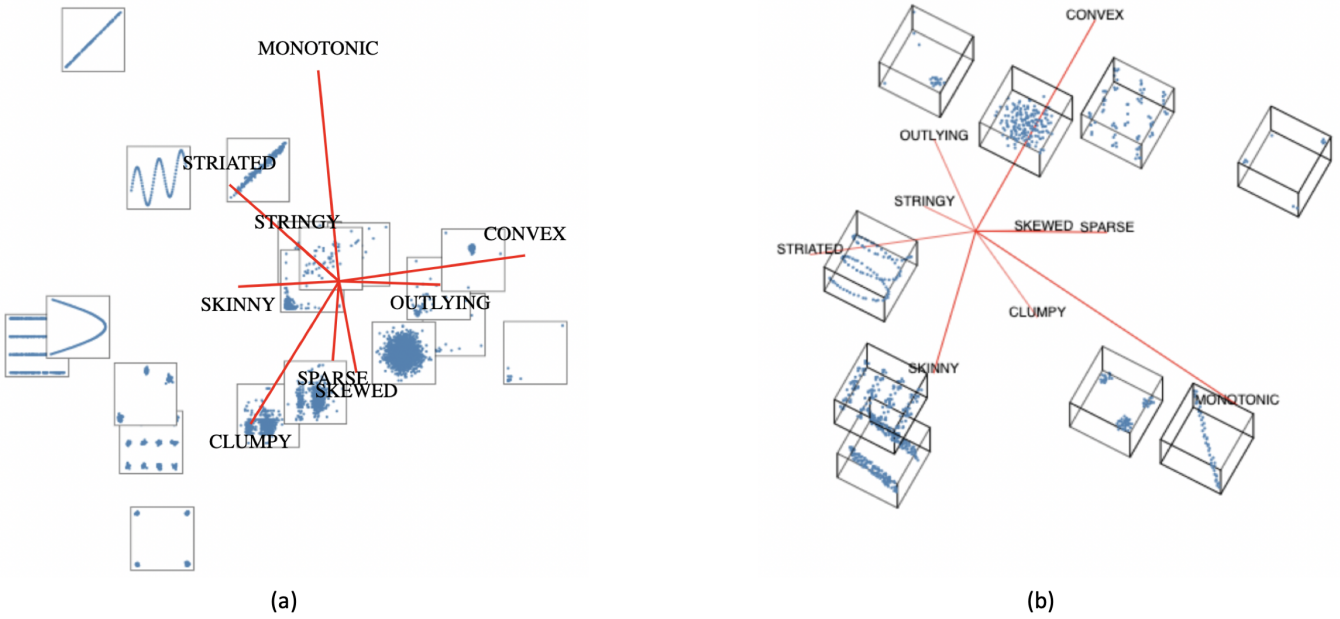

Figure 2: Scagnostics measures and their targeting patterns projected using two and three PCA principal components in panel (a) and panel (b) correspondingly.

Figure 2 graphically depicts the Scagnostics measures projected into 2D and 3D space using principal component analysis (PCA) [WEG87]. Panel (a) is the biplot [Gab71] of the measures and their exemplar scatterplots. In this case, the nine measures are projected in 2D space using the first two principal components. The red lines are the projected unit vectors for the nine corresponding Scagnostics measures. The scatter patterns are placed at their 2D projected locations of their Scagnostics measures. Similarly, for the 3D case in Panel (b), the first three principal components are examined. It is observable that the scatters are placed near to their corresponding unit vectors of the measures that they are designed to flag.

2. Scagnostics examples

To show the usefulness of ScagnosticsJS in practice, we applied it (all 2D, 3D, and nD implementations) in a web interface to monitor the computer health status of an High Performance Computing (HPC) center at a university [BK17]. It shows that the library is capable of performing its tasks in real-time [Dan19]. Figure 3 shows examples of multivariate outliers detected via our web monitoring interface using the nD version. In particular, one outlier (computer) on the right panel shows a burning CPU2 temperature (100°F) and very high fan speeds (at the bottom) which had been detected and replaced before harming the entire system. Radar chart [MMP09] was selected instead of parallel coordinates since it can capture the morphology of multidimensional curves [Saa08].

3. Scagnostics exploration page

The source codes and ScagnosticsJS API reference page are available at `https://idatavisualizationlab.github.io/ScagnosticsJS`. In addition, interested readers can explore the stages of Scagnostics calculations as well as how the Scagnostics measures reacts to their target patterns in scatterplots from the following exploration pages:

- 2D: `https://idatavisualizationlab.github.io/ScagnosticsJS/scagnostics`
- 3D: `https://idatavisualizationlab.github.io/ScagnosticsJS/scagnostics3d`
- nD: `https://idatavisualizationlab.github.io/ScagnosticsJS/scagnosticsnd`

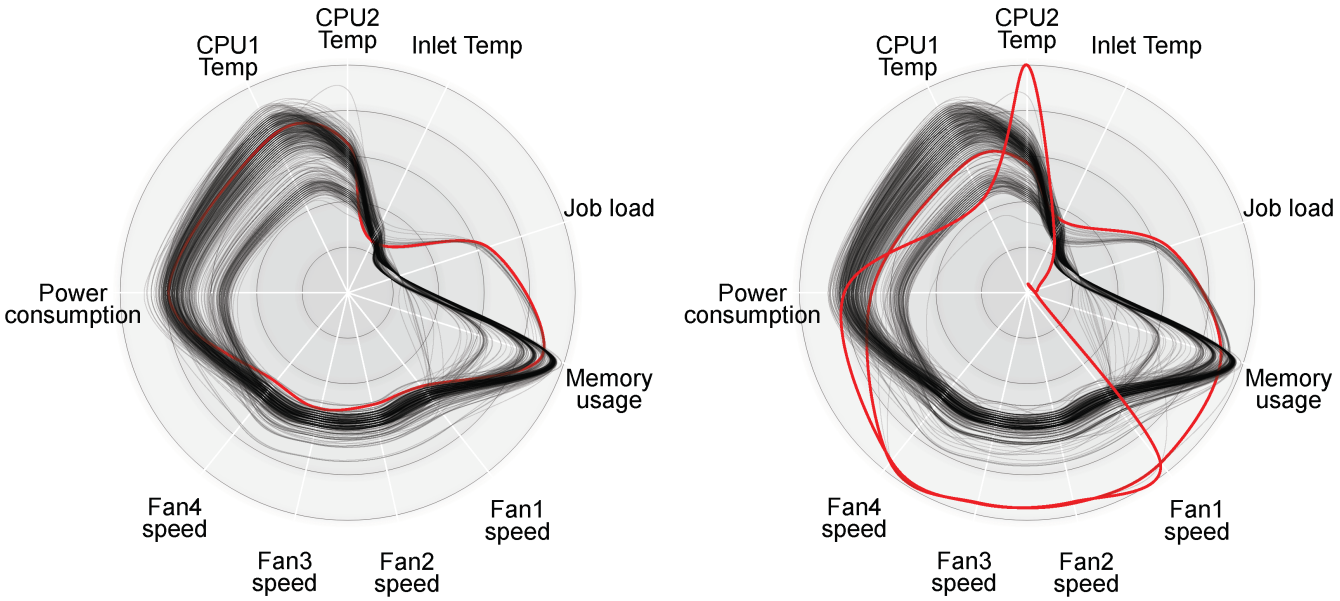

Figure 3: Red curves show suspicious computers in a HPC center.

References

[BK17] BETKE, EUGEN and KUNKEL, JULIAN. "Real-time I/O-monitoring of HPC applications with SIOX, elasticsearch, Grafana and FUSE". *International Conference on High Performance Computing*. Springer. 2017, 174–186 2.

[Dan19] DANG, TOMMY. "Visualizing Multidimensional Health Status of Data Centers". *Programming and Performance Visualization Tools*. Ed. by BHATELE, ABHINAV, BOEHME, DAVID, LEVINE, JOSHUA A., et al. Cham: Springer International Publishing, 2019, 273–283. ISBN: 978-3-030-17872-7 2.

[Gab71] GABRIEL, KARL RUBEN. "The biplot graphic display of matrices with application to principal component analysis". *Biometrika* 58.3 (1971), 453–467 2.

[MMP09] MEYER, MIRIAH, MUNZNER, TAMARA, and PFISTER, HANSPETER. "MizBee: A Multiscale Synteny Browser". *IEEE Transactions on Visualization and Computer Graphics* 15.6 (Nov. 2009), 897–904. ISSN: 1077-2626. DOI: 10.1109/TVCG.2009.167. URL: http://dx.doi.org/10.1109/TVCG.2009.167 2.

[Saa08] SAARY, M JOAN. "Radar plots: a useful way for presenting multivariate health care data". *Journal of clinical epidemiology* 61.4 (2008), 311–317 2.

[WEG87] WOLD, SVANTE, ESBENSEN, KIM, and GELADI, PAUL. "Principal component analysis". *Chemometrics and intelligent laboratory systems* 2.1-3 (1987), 37–52 2.