



City Research Online

City St George's, University of London

Citation: Wood, J. (2016). Visual Analytic Design for Contextualising Sensor Data. Paper presented at the VIS 2016, 23-28 Oct 2016, Baltimore, USA.

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/15523/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Colour hue is used to symbolise the period spent in any given location with bars of higher saturation symbolising prox card entry/exit

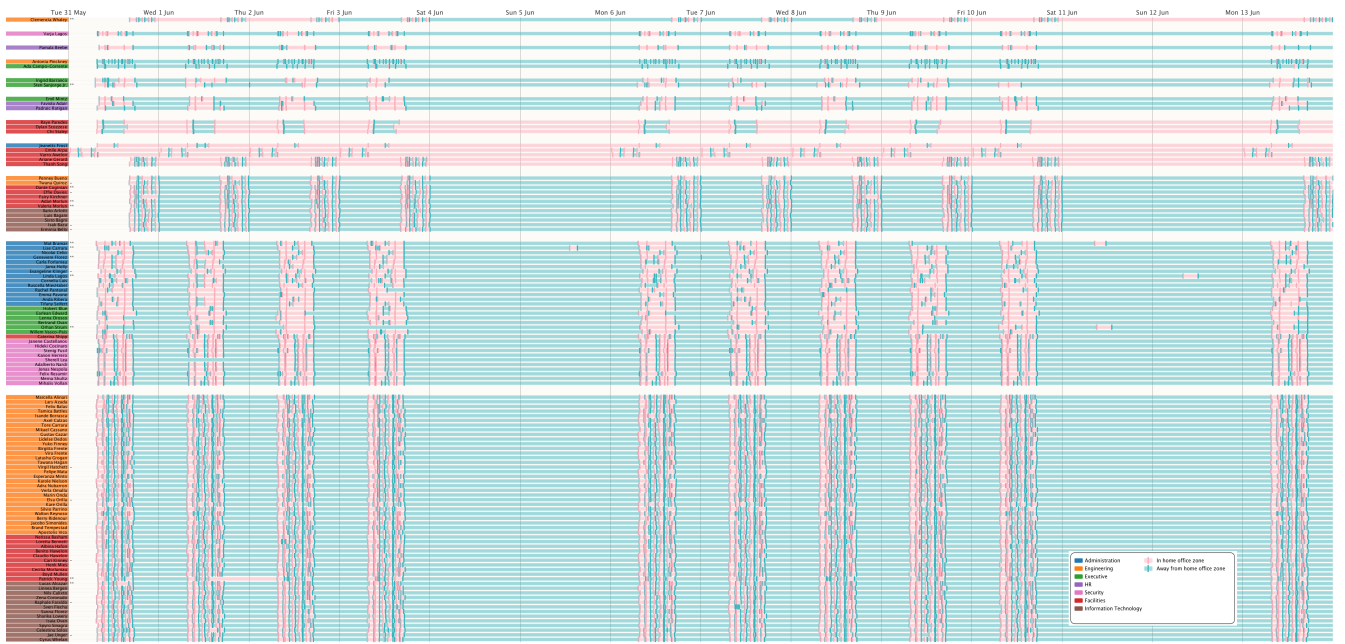


Figure 2: Employee view showing employees' location in their 'home' office zone (pink) or outside their home zone (blue) organised vertically by cluster and horizontally by time.

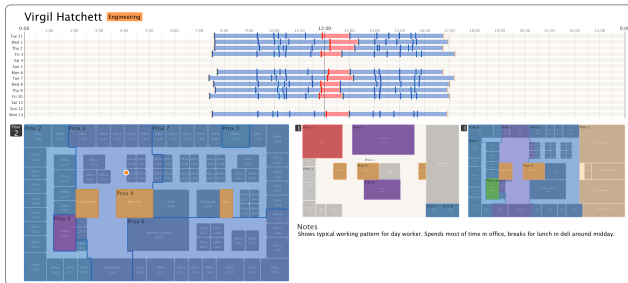


Figure 3: Single employee view showing calendar-type layout and classified locations.

points and lower saturation for interpolated location between prox card readings. While the prox card zone (Figure 1) can be mapped to hue, this provides little direct indication of the likely activity of each employee. So zones were additionally classified by their dominant function (e.g. office space, conference facility, dining etc.) and each given a unique hue evenly spaced in perceptual colour space. Where zones comprised more than one function (e.g. combined office and conference space), colours were linearly interpolated between classes to give some visual indication of classification uncertainty. Finally a zones were classified as being either a 'home' location in which an employee's regular office was located, or an 'away' location in a different prox zone (see Figure 2).

2.3 Cluster-driven Layout

To support both characterisation of routine behaviour and anomalous departure from routine, every employee's two-week pattern of zone location, home-away location and location function was represented as a multidimensional vector and clustered using multiple runs of k-means++ [1]. Despite the relative stability of this form of clustering, large vectors can produce different clusters when run on the same dataset so interaction permitted both re-runs of the clustering and dynamic changes in the number of clusters in order to vi-

ually assess their stability. By ordering employees by cluster size, common and anomalous patterns were separated visually (common patterns towards the bottom, unique patterns towards the top).

3 CONCLUSION

Analysis of the imprecise geosensor 'checkpoint' network data and comparison between staff movement and building environment was facilitated by a set of simple design choices that (i) reflected the importance of change over time in the inference of critical behaviour; (ii) used clustering and layout to automatically separate routine from anomalous movement patterns; (iii) recognised the uncertainty in classification through both symbolisation and the role of interaction. An interactive visual analytic system built following these design elements provided a means of exploring patterns; hypothesising causal and associative relationships within the data set; recording the provenance of observed patterns and inferences [4]; and communicating findings by the analysts.

REFERENCES

- [1] D. Arthur and S. Vassilvitskii. k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1027–1035. Society for Industrial and Applied Mathematics, 2007.
- [2] M. Duckham. *Decentralized spatial computing: foundations of geosensor networks*. Springer Science & Business Media, 2012.
- [3] M. Duckham, M. van Kreveld, R. Purves, B. Speckmann, Y. Tao, K. Verbeek, and J. Wood. Modeling checkpoint-based movement with the earth mover's distance. *Lecture Notes in Computer Science*, 9927, 2016.
- [4] R. Walker, A. Slingsby, J. Dykes, K. Xu, J. Wood, P. H. Nguyen, D. Stephens, B. W. Wong, and Y. Zheng. An extensible framework for provenance in human terrain visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2139–2148, 2013.
- [5] J. J. V. Wijk. Cluster and calendar based visualization of time series data. In *INFOVIS '99: Proceedings of the 1999 IEEE Symposium on Information Visualization*, pages 4–9. IEEE Computer Society, 1999.
- [6] J. Wood. Visualizing personal progress in participatory sports cycling events. *IEEE Computer Graphics and Applications*, 35(4):73–81, 2015.