Title	Sankey Network Diagrams to Depict Bulk Power Transactions for Operator Situational Awareness
Authors(s)	Cuffe, Paul, Beiranvand, Arash
Publication date	2021-07-02
Publication information	Cuffe, Paul, and Arash Beiranvand. “Sankey Network Diagrams to Depict Bulk Power Transactions for Operator Situational Awareness.” IEEE, July 2, 2021. https://doi.org/10.1109/PowerTech46648.2021.9494795.
Conference details	PowerTech 2021, Madrid, Spain (held online due to Coronavirus outbreak), 28 June - 2 July 2021
Publisher	IEEE
Item record/more information	http://hdl.handle.net/10197/25713
Publisher's statement	© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
Publisher's version (DOI)	10.1109/PowerTech46648.2021.9494795

Sankey Network Diagrams to Depict Bulk Power Transactions for Operator Situational Awareness

1st Arash Beiranvand
School of Electrical and Electronic Engineering
University College Dublin
Dublin, Ireland
arash.beiranvand@ucdconnect.ie

2nd Paul Cuffe
School of Electrical and Electronic Engineering
University College Dublin
Dublin, Ireland
paul.cuffe@ucd.ie

***Abstract*—It is not always easy to apprehend the prevailing disposition of power flows through an electrical network. Typically, power system one line diagrams are drawn in a pseudo geographic way. This depiction may obscure where the bulk power flows or the corridors of coherent flow are for a particular operational snapshot. For instance, there could be two electrical regions exchanging large flows of power, yet a geographic layout could obscure this. Bulk power flows are of interest as they are intimately connected with a systems operational security. In this paper a new technique is proposed to more meaningfully portray the coherent flows of active power through a transmission system, based on the Sankey diagram. Furthermore, relying on the *Line Outage Distribution Factors*, a novel line-oriented clustering methodology is applied to the diagram. This novel clustering provides useful insights regarding the possible contingencies within the grid and the likely post-outage power flow conditions.**

I. INTRODUCTION

In this paper, a novel flow-oriented visualisation approach using the Sankey diagram [1]–[3] is proposed to depict power flow dispositions within power grids. The suggested visualisation technique is proposed to make it easier to trace patterns of heavy flow through a grid. Such heavy coherent flow, and associated congestion bottlnecks, are of interest as any interruptions in their functions can lead to serious operational vulnerabilities like cascading outages and consequently, load-shedding and frequency problems [4]–[6].

The Sankey diagram is often used to show the efficiencies of thermodynamic cycles or patterns of inputs and outputs in integrated energy systems. In this paper, we apply the Sankey diagram directly to portray the physical powers flows within a power grid. Lines are oriented such that all power flows are from the top to the bottom of the diagram, and the thickness of each line shows the magnitude of its active power flow. This allows a rapid identification of the most important bulk power exchanges that arise for a particular generation dispatch snapshot.

The proposed flow-oriented visualisation technique is enhanced further by applying a novel line-oriented clustering methodology. The clustering technique uses the data extracted from the *Line Outage Distribution Factors* (LODF) matrix[7]. The LODF matrix measures how a line's pre-outage power flow will be redistributed onto adjacent lines if it is tripped off [8]. Hence, we rely on it to investigate what alternative path(s) would be available for each line's flow if it is interrupted. In this way, the Sankey diagram of the visualised power grid is sectioned into a certain number of zones, each of which shows in the case of removing any of the constituent lines where its flow mostly will be redistributed. A significant advantage of this feature is that, for instance, the operators could detect what alternative paths are available for a heavy flow within the grid and, considering the pre-outage flows of the alternative paths, whether the alternative paths are capable of handling the redistributed flow. Thus, visualising the grids by our novel technique, the operators, first, are more aware of coherent flow paths and congestion bottlenecks in the transmission system, secondly, have insights into the post-contingency condition of the grid and the possible vulnerabilities.

Concerning networks visualisation, the geographically-oriented techniques usually use single line diagrams to model the grids [9]–[13]. Different approaches are followed in this type of visualisation to enhance the basic single line diagram. For example, in [9] and [10], the voltage magnitudes are visualised using coloured contour lines. The author in [13] adds a third-dimension to the diagram and argues the three-dimensional visualisation of power networks can serve for educational purposes to analysing complex concepts.

The connectivity-oriented visualisation, also known as the structural-oriented visualisation, approaches are another way of diagramming an electrical grid. Work in [14], which is based on the connectivity-oriented visualisation, has demonstrated that power system diagrams which explicitly show inter-node electrical distances allow the rapid identification of electrically remote buses.

Conventional single line diagrams typically show branch power flows radiating out from buses in diverse and disorganised directions [15]. This impedes the ability of an analyst to trace the paths of active power from generators to loads, even if arrows or animations are used to denote flow directions, as

---

A. Beiranvand (arash.beiranvand@ucdconnect.ie) and P. Cuffe are with the School of Electrical and Electronic Engineering, University College Dublin. This publication has emanated from research supported in part by Science Foundation Ireland (SFI) under the SFI Strategic Partnership Programme Grant Number SFI/15/SPP/E3125 and additional funding provided by the UCD Energy Institute. The opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Science Foundation Ireland. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission

978-1-6654-3597-0/21/$31.00 ©2021 IEEE

in [16] or [17].

The flow-oriented visualisation approaches are a more recent grid visualisation approach that seeks to more explicitly depict the power flow dispositions within the power grids [5], [6], [18], [19]. Among the available flow-oriented visualisation techniques, the methodologies presented in [5] and [6] particularly focus on detecting and visualising the coherent flow-gates within the grids. To this end, power grids are modelled as the Directed Acyclic Graphs (DAG) [20] so that the lines' power flow directions set the directionality of the edges. In [6], using the topological sorts of the obtained DAG for a power grid, the coherent flow-gates within the grid are detected. In [5], it is argued that, if the function of a power system is conceptualised as delivering the power from the sources to the sinks within the DAG, some nodes are working prominently to facilitate these power transactions. Such sets of nodes form 'the corridors of coherent flow' within the grid [5]. Then, the Sugiyama layout [21] is used for the power grids' DAG to visualise and explicitly portray these corridors of coherent flows.

Heavy flow-gates in power grids [4] in general, the heavy coherent flow-gates in particular [6], are reported as some of the most important operational vulnerabilities for power systems where significant amounts of power flows are exchanged between different regions within the grids. It is shown in [4] and [6] that outages within such flow-gates may notably damage the grids, in terms of the load-shedding and frequency issues. Hence, appropriately visualising the power flow dispositions can help to detect bulk power flow transmissions and consequently, boost the situational awareness of the control room.

The proposed visualisation technique in the present paper is a pure flow-oriented visualisation approach where we are not interested in the nodes' geographical locations, but instead position buses to portray their role in the bulk power transactions arising in a particular generation dispatch snapshot. The proposed visualisation and clustering methodology is a step towards filling this gap in the power system visualisation literature. The proposed Sankey diagram can be considered as an extension to the Sugiyama layout for power grids in [5].

The novel diagramming methodology is described in section II, some examples and discussion are given in section III, while section IV concludes the work.

II. METHODOLOGY

In this section, the novel visualisation technique is presented. Firstly, the data obtained from the power flow calculations are used to set the direction and value of every flow within the grid. Then, the Sankey diagram is portrayed. Finally, the proposed clustering methodology is described and applied to partition the diagram.

A. Flow-oriented diagramming

The lines' power flows could be obtained through various methodologies such as simulation, state estimation or direct metering. Notably, both the grid's static topology and the load-generation profile affect the power flows.

Considering the DC-PF power flow assumptions, the flow of line l, connecting buses n and m, is calculated as follows:

$$P_{n,m} = \frac{\delta_n - \delta_m}{X_l} \tag{1}$$

Where, δ_n and δ_m are the voltage angles of buses n and m and X_l is the reactance of l.

Equation (1) determines both the magnitude and directionality of every line's flow. Concerning the flow directionality, a positive value of $P_{n,m}$ in (1) yields $n \rightarrow m$ direction for the edge between nodes n and m. Consequently, obtaining the results from (1) for the grid, the Sankey diagram of the network digraph could be portrayed.

B. Clustering the lines within the grid

1) Line Outage Distribution Factors: The Line Outage Distribution Factors (LODF) matrix is used to cluster the lines within the grid into functional groups. This builds on the idea of designating 'tributary' and 'distributary' type corridors outlined in [5], which borrows the terminology for categorising river systems [22].

To calculate the LODF matrix, the Power Transfer Distribution Factor matrix, Π [8] should be obtained for the grid. Under DC power flow approximations, the component at row l and column t within Π, ${\Pi_l}^{st}$, shows the flow change in line l, connecting node n to node m, due to the injection of 1 MW at bus s and withdrawal at bus t and can be calculated as follows [8]:

$${\Pi_l}^{st} = \frac{(x_{ns} - x_{nt}) - (x_{ms} - x_{mt})}{X_l} \tag{2}$$

Where x_{is} is the imaginary part of the component at row i and column s within the grid's impedance matrix, Z_{bus}. Also, Z_{bus} is the inverse of the grid's admittance matrix, Y_{bus}.

In [8], it is proven that the corresponding component of l within LODF matrix for removing a line k , connecting buses i and j, is calculated as follows:

$${\zeta_l}^{k} = \frac{{\Pi_l}^{ij}}{1 - {\Pi_k}^{ij}} \tag{3}$$

Using (3), the percentage of the pre-outage flow in line k which will be redistributed onto line l after removing line k is calculated. We rely on this important feature of the LODF matrix to detect the alternative path(s) available for the flow in every line of the grid. Applying a higher level clustering technique for ζ matrix enables us to group the lines into a certain number of clusters. Each group shows where the power flows of its constituent lines would mostly be redistributed if any of them is removed.

2) Applying the clustering technique: For using ζ matrix to cluster the lines, some initial steps need to be taken to prepare it. Firstly, the radial lines of the grids need to be excluded from the targeted lines for clustering. This is because a radial line removal would never affect the other lines' flows and therefore, the radial lines do not show meaningful correlations with the other lines of the grid in this sense. Consequently, at the first step, the corresponding rows and lines of the radial lines within ζ are removed, giving the modified LODF matrix, ζ'.

At the next step, the dissimilarity matrix [23], D, is obtained for ζ' as follows:

$$D = I - \zeta' \tag{4}$$

where I is the *all-ones matrix* [24] whose dimension is the same as ζ'. Finally, the diagonal components of D are artificially set to zeros. The resulting matrix is a symmetric matrix with all diagonal components equal to zero.

Finally, to cluster the grid, the *K-medoids* clustering algorithm [25] is applied to D. Each row in D represents a line of the grid. Consequently, the *K-medoids* clustering algorithm considers each row of D as the data related to the corresponding line and by this way, clusters the lines into a certain number of groups.

III. Results

Two sample grids from the repository at [26] are selected for this study. The parallel lines within each grid are merged and replaced with their equivalent lines. The power flow calculations and clustering methodology are handled by MATPOWER [27]. Also, the Sankey diagrams are portrayed using the Plotly library [28] on Jupyter Notebook [29]. The raw data and scripts created for the suggested methodology could be found in [30].

Concerning the clustering, we cluster the lines within each sample grid into 5 groups. One may put the lines into more or fewer clusters. The more clusters the lines are grouped in, the more engaged lines within each cluster to ease the power transactions.Also, the color sets used in this paper exclusively aim to distinguish the clusters and are not proportional to the lines' loading. After clustering the non-radial lines, different approaches could be followed to come up with the radial lines, as they have already been excluded for clustering. In this paper, each radial line is put into the corresponding cluster of the non-radial line with the biggest flow connected to either side of the radial line. If all the connecting lines to both sides are radial, the biggest flow which has already been clustered is selected. Otherwise, the biggest flow is chosen.

A. *Visualising* `pglib_opf_case39_epri`

Fig. 1 shows the obtained layout for `pglib_opf_case39_epri` using the proposed visualisation technique. We have considered vertical orientation, the same as the Sugiyama layout presented in [5], for the Sankey diagram which tangibly portrays how the active power flows from the top to the bottom of the diagram. Line colours denote cluster membership. The radial lines in each cluster, if any, are distinguished from the other lines within the cluster by more transparent colour.

This visualisation reveals important details regarding the power flow dispositions within the network. For instance, Fig. 1 shows node 31 is an important source node [5] for the grid which is injecting a significant amount of power into the network. Using the obtained layout, the operators can easily notice an important generator is connected to this node. Also, as an important feature of the proposed Sankey diagram, just by looking at a node, the operators can detect if a big load is connected to the node or, vice versa, the node functions as a generation node within the grid. To this end, it is just sufficient to compare the thickness of the incoming and outgoing flows. In the case of connecting a big load to the node, the outgoing flows are notably thinner, in total, than the incoming flows, and vice versa for a generation node. For example, Fig. 1 shows a big load is connected to node 4. The diagram also shows that nodes 33 and 34 are mainly responsible for supplying 19 and 20, with limited interaction with the wider network.

This novel visualisation technique effectively portrays the power flow dispositions within the grid which results in significantly improving the situational awareness of the control room.

The clustering methodology also brings further useful insights regarding the operational vulnerabilities to the operators. For example, consider the heavy flow in the line connecting nodes 6 and 5 in Fig. 1. The clustering technique shows the other lines within the corresponding cluster of this line, coloured in green, are mostly handling heavy flows. Consequently, in the case of removing the line connecting nodes 6 and 5, its large flow will mostly be redistributed onto the other adjacent lines within the cluster whose pre-outage flows are significant. These facts suggest the line and specific area could be important operational vulnerabilities for the grid.

B. *Visualising* `pglib_opf_case118_ieee`

Fig. 2 shows the obtained layout for `pglib_opf_case118_ieee` using the proposed visualisation methodology. For this larger system, the diagram is noticeably more cluttered, with many of the lines overlapping each other. This somewhat restricts the legibility of the diagram, and suggests that bespoke node layout algorithms may be necessary to produce scalable network visualisations of this type. Alternatively, the smaller flows and lower voltage levels could be excluded, to produce a summary diagram of the principal bulk power transactions.

Notwithstanding these limitations, Fig. 2 shows that node 69 is a very important source node where a significant amount of power is injected into the grid and also, many other nodes are being supplied by this node. Such an important node is easily detected by our suggested visualisation technique. A node such as 49 is seen as an important nexus, where three power inflows are split to six outflows. Also, two major sinks, 90 and 116, can be detected easily within the grid which are notable load nodes. Also, by comparing the incoming and outgoing flows to the nodes, some load and generation busses can be detected. For example, Fig. 2 shows notable load are connected to nodes 77 and 62 and 82. As it is obvious, the power flow dispositions can be traced effectively by our novel visualisation methodology.

The implemented clustering technique shows a specific section of the grid coloured in purple may be a reasonably secure part of the grid as it includes many well-connected lines with light flows. These features suggest a contingency in that section is likely to be tolerated by the grid. These functional insights regarding the operation state can notably boost the situational awareness of the control room.

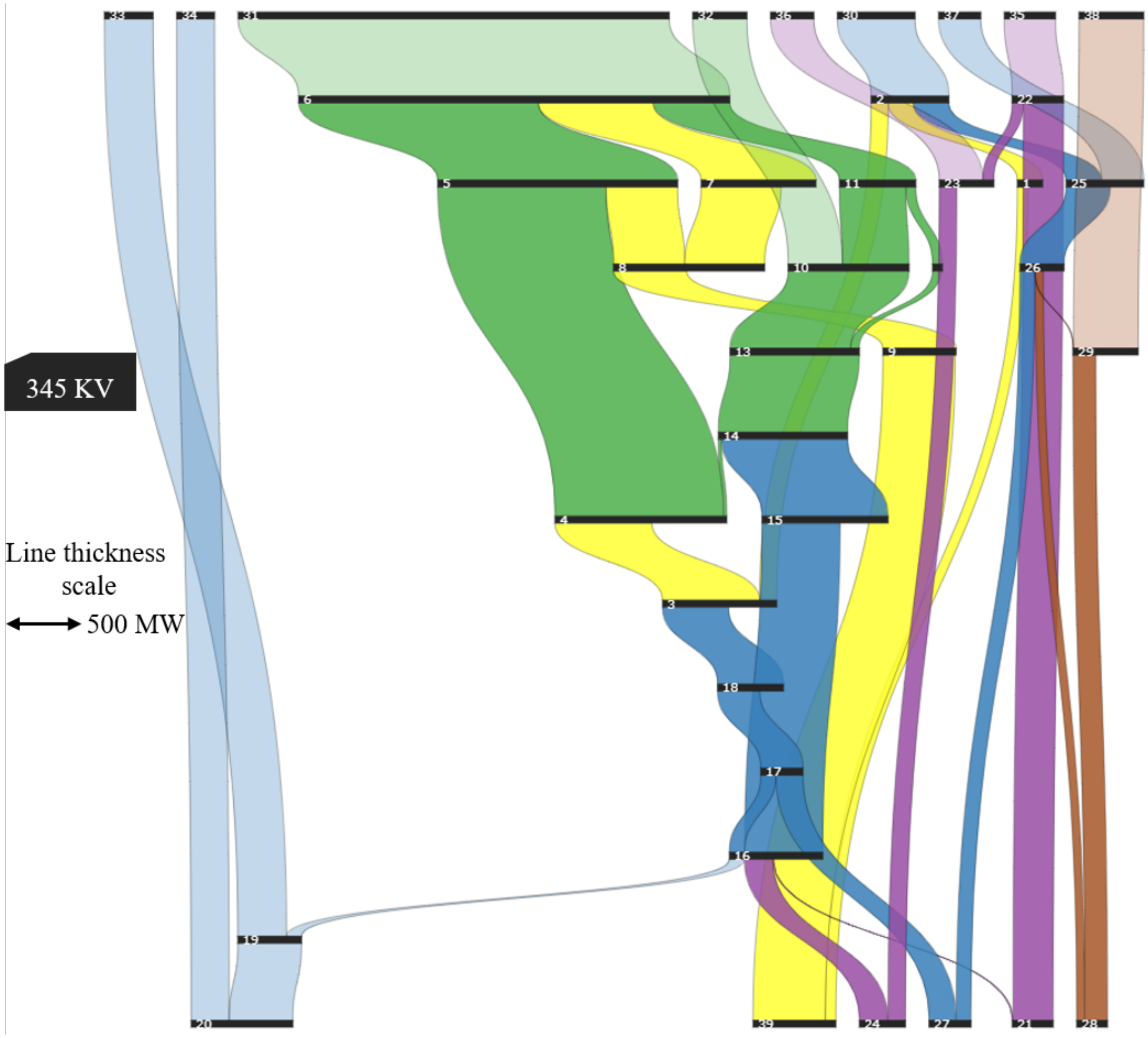

Fig. 1. The Sankey diagram for `pglib_opf_case39_epri` Active power flows from the sources at the top to the sinks at the bottom of the diagram, and the magnitude of such flows is represented by their line thickness. The horizontal rectangles are buses, whose nominal voltage is encoded by their shade of grey.

C. Discussion

The novel proposed visualisation technique in this paper can significantly enhance the situational awareness of the control room. That is, as shown by providing a couple of examples in Section III, the visualisation is capable of spotting significant generation nodes directly. Likewise, by comparing the thicknesses of the incoming and outgoing flows to the nodes, notable loads can be identified readily. Also, the applied clustering technique further enhances the situational awareness of the control room by identifying the sets of branches that are closely meshed to ease the power transactions within the grid. This clustering technique reveals where the flow of a particular line will be mostly redistributed after its removal. In case that the alternative paths for the line's flow are already handling heavy loading, the line and its alternative paths can be potential operational vulnerabilities for the grid. All of these vital practical details regarding the operation condition are directly revealed by the suggested visualisation technique.

Finally, comparing the visualisations obtained from implementing the proposed technique for `pglib_opf_case39_epri` and `pglib_opf_case118_ieee` suggests for bigger power grids the obtained layouts may be cluttered. To avoid cluttering the portrayals for the bigger grids, one could apply the network's size reduction methodologies like the technique proposed in [5] to efficiently reduce the network's size and consequently obtain more explicit layouts of the grids.

IV. CONCLUSIONS

In this paper, a novel flow-oriented visualisation technique, using Sankey diagrams, was proposed to explicitly depict the

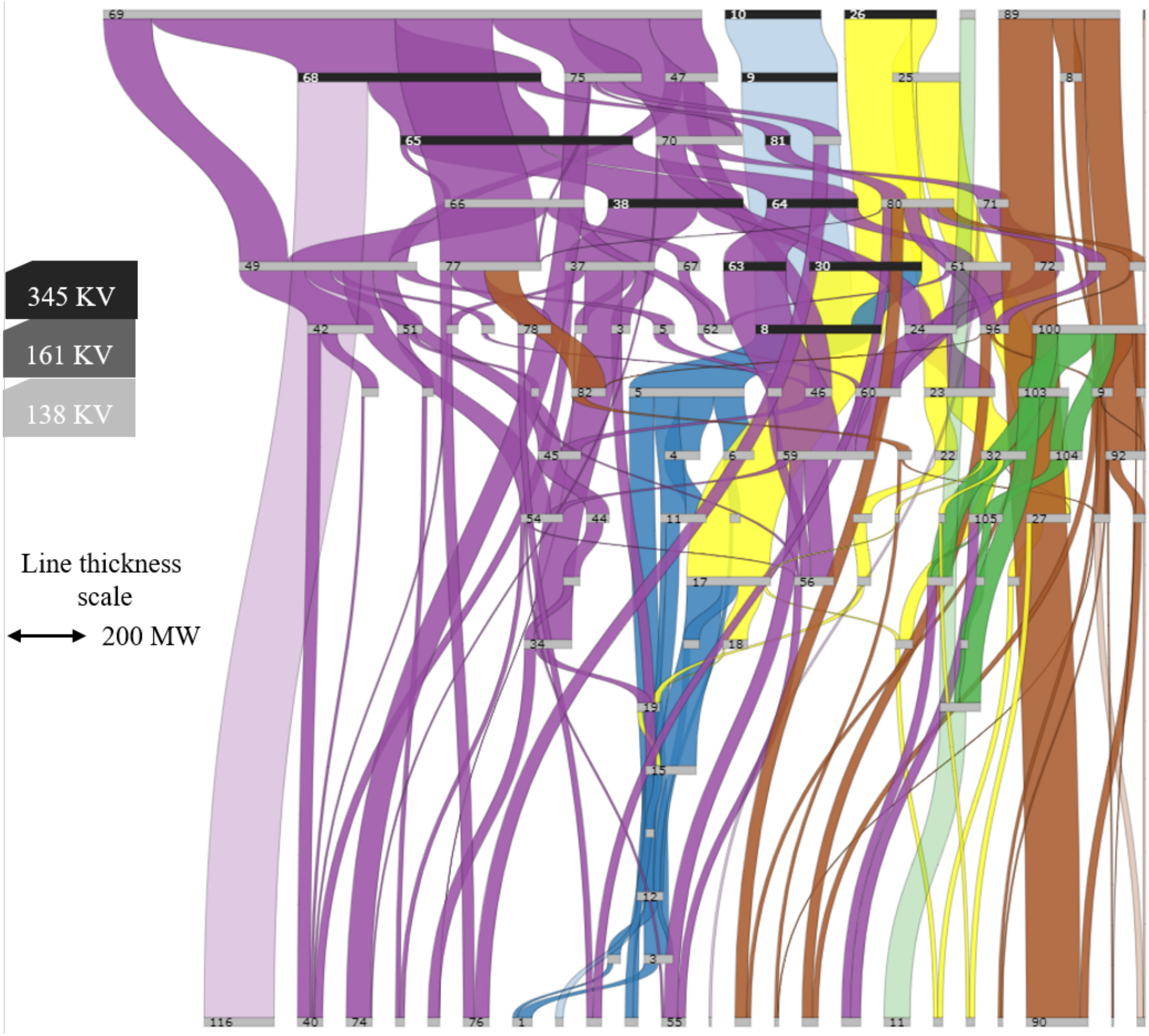

Fig. 2. The Sankey diagram for `pglib_opf_case118_ieee`. Active power flows from the sources at the top to the sinks at the bottom of the diagram, and the magnitude of such flows is represented by their line thickness. The horizontal rectangles are buses, whose nominal voltage is encoded by their shade of grey.

disposition of bulk power flows. A novel line-oriented clustering methodology was also applied to the obtained diagram as a step towards providing operators with useful insights regarding potential post-contingency power flow condition of the grid. Applying the suggested visualisation technique for two sample grids showed it effectively enables us to trace the power flow dispositions within the grid which results in significantly boosting the situational awareness of the control room. However, for the larger network diagram there were problems with legibility and visual clutter. Future work may consider new algorithmic approaches to positioning nodes to produce cleaner power system flow diagrams of this type.

REFERENCES

[1] P. Riehmann, M. Hanfler, and B. Froehlich, “Interactive sankey diagrams,” in *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.*, 2005, pp. 233–240.

[2] C. Deng, H. Li, and Y. Shao, “Research on the drawing method of energy sankey diagram based on java,” in *16th International Conference on Advanced Communication Technology*, 2014, pp. 88–91.

[3] D. C. Zarate, P. L. Bodic, T. Dwyer, *et al.*, “Optimal sankey diagrams via integer programming,” in *2018 IEEE Pacific Visualization Symposium (PacificVis)*, 2018, pp. 135–139.

[4] R. Sen Biswas, A. Pal, T. Werho, *et al.*, “A graph theoretic approach to power system vulnerability identification,” *IEEE Transactions on Power Systems*, pp. 1–1, 2020.

[5] J. J. John, A. Beiranvand, and P. Cuffe, “Clustering nodes in a directed acyclic graph by identifying corridors of coherent

flow," in *2020 6th IEEE International Energy Conference (ENERGYCon)*, 2020, pp. 604–609.

[6] A. Beiranvand and P. Cuffe, "A topological sorting approach to identify coherent cut-sets within power grids," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 721–730, 2020.

[7] Á. S. Xavier, F. Qiu, F. Wang, *et al.*, "Transmission constraint filtering in large-scale security-constrained unit commitment," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 2457–2460, 2019.

[8] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, in *Power generation, operation, and control*, Wiley, 2014.

[9] J. D. Weber and T. J. Overbye, "Voltage contours for power system visualization," *IEEE Transactions on Power Systems*, vol. 15, no. 1, pp. 404–409, 2000.

[10] T. J. Overbye, D. A. Wiegmann, A. M. Rich, *et al.*, "Human factors aspects of power system voltage contour visualizations," *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 76–82, 2003.

[11] Yan Sun and T. J. Overbye, "Visualizations for power system contingency analysis data," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1859–1866, 2004.

[12] M. J. Laufenberg, "Visualization approaches integrating real-time market data," in *IEEE PES Power Systems Conference and Exposition, 2004.*, 2004, 1550–1555 vol.3.

[13] F. Milano, "Three-dimensional visualization and animation for power systems analysis," *Electric Power Systems Research*, vol. 79, no. 12, pp. 1638–1647, 2009, ISSN: 0378-7796.

[14] P. Cuffe and A. Keane, "Visualizing the electrical structure of power systems," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1810–1821, 2017.

[15] P. Cuffe and A. Keane, "Novel quality metrics for power system diagrams," in *2016 IEEE International Energy Conference (ENERGYCON)*, 2016, pp. 1–5. DOI: 10.1109/ENERGYCON.2016.7514086.

[16] D. A. Wiegmann, G. R. Essenberg, T. J. Overbye, *et al.*, "Human factor aspects of power system flow animation," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1233–1240, 2005.

[17] T. J. Overbye and J. D. Weber, "New methods for the visualization of electric power system information," in *IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings*, 2000, pp. 131–16c.

[18] P. Kumar and A. K. Singh, "Transmission line power flow visualization using 3-dimensional circle diagrams," in *2015 Annual IEEE India Conference (INDICON)*, 2015, pp. 1–6.

[19] S. Sugino, K. Okabe, N. Komuro, *et al.*, "Power-flow simulation with visualization function based on ieee common data format," in *2016 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, 2016, pp. 301–304.

[20] R. A. Sahner and K. S. Trivedi, "Performance and reliability analysis using directed acyclic graphs," *IEEE Transactions on Software Engineering*, vol. SE-13, no. 10, pp. 1105–1114, 1987.

[21] K. Sugiyama, S. Tagawa, and M. Toda, "Methods for visual understanding of hierarchical system structures," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 11, no. 2, pp. 109–125, Feb. 1981.

[22] C. R. Fielding, P. J. Ashworth, J. L. Best, *et al.*, "Tributary, distributary and other fluvial patterns: What really represents the norm in the continental rock record?" *Sedimentary Geology*, vol. 261-262, pp. 15–32, 2012, ISSN: 0037-0738. DOI: https://doi.org/10.1016/j.sedgeo.2012.03.004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0037073812000723.

[23] I. Amerise and A. Tarsitano, "Combining dissimilarity matrices by using rank correlations.," *Computational Statistics*, vol. 31, no. 1, pp. 353–367, 2016.

[24] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd. USA: Cambridge University Press, 2012, ISBN: 0521548233.

[25] A. Struyf, M. Hubert, and P. Rousseeuw, "Clustering in an object-oriented environment," *Journal of Statistical Software, Articles*, vol. 1, no. 4, pp. 1–30, 1997.

[26] S. Babaeinejadsarookolaee, A. Birchfield, R. D. Christie, *et al.*, "The power grid library for benchmarking ac optimal power flow algorithms," 2019. arXiv: 1908.02788 `[math.OC]`.

[27] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

[28] P. T. Inc. (2015). "Collaborative data science."

[29] T. Kluyver, B. Ragan-Kelley, F. Pérez, *et al.*, "Jupyter notebooks-a publishing format for reproducible computational workflows," in *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, F. Loizides and B. Scmidt, Eds., IOS Press, 2016, pp. 87–90.

[30] A. Beiranvand and P. Cuffe, "Raw data and scripts from Towards a Flow-oriented Visualisation Approach for Electric Power System Situational Awareness," [Online]. Available: https://figshare.com/s/e17e8080d6de8ad29cc5.