EUROGRAPHICS 2018 / J. Heer, H. Leitte, and T. Ropinski
(Guest Editors)

Volume 37 (2018), Number 3

Hunting High and Low: Visualising Shifting Correlations in Financial Markets

P. M. Simon[1,2] and C. Turkay[1]

[1]Department of Computer Science, City, University of London, UK
[2]Scaridae Analytics, London, UK

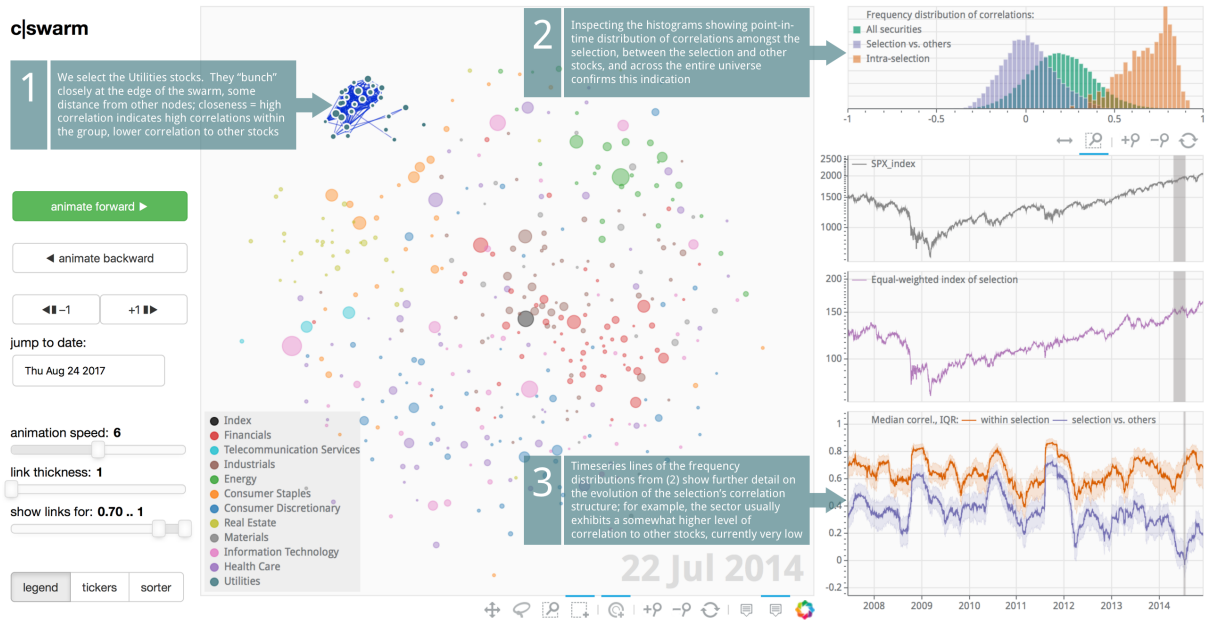

Figure 1: *The c|swarm user interface, showing S&P 500 (US large capitalisation) stocks' correlations of daily returns for the three months to 22 July, 2014. The utilities industry sector is selected; the analysis workflow suggests that these stocks could be useful diversifiers.*

Abstract

The analysis of financial assets' correlations is fundamental to many aspects of finance theory and practice, especially modern portfolio theory and the study of risk. In order to manage investment risk, in-depth analysis of changing correlations is needed, with both high and low correlations between financial assets (and groups thereof) important to identify. In this paper, we propose a visual analytics framework for the interactive analysis of relations and structures in dynamic, high-dimensional correlation data. We conduct a series of interviews and review the financial correlation analysis literature to guide our design. Our solution combines concepts from multi-dimensional scaling, weighted complete graphs and threshold networks to present interactive, animated displays which use proximity as a visual metaphor for correlation and animation stability to encode correlation stability. We devise interaction techniques coupled with context-sensitive auxiliary views to support the analysis of subsets of correlation networks. As part of our contribution, we also present behaviour profiles to help guide future users of our approach. We evaluate our approach by checking the validity of the layouts produced, presenting a number of analysis stories, and through a user study. We observe that our solutions help unravel complex behaviours and resonate well with study participants in addressing their needs in the context of correlation analysis in finance.

CCS Concepts

•***Human-centered computing*** → ***Visual analytics;***

1. Introduction

Correlation analysis is fundamental to many aspects of finance theory, especially portfolio theory and the study of financial risk. In his seminal paper, which founded modern mean-variance portfolio theory sixty-five years ago, Markowitz [Mar52] discussed the importance of diversification in investment portfolio construction—the spreading of an investor's money across multiple diverse securities in order to reduce risk (an investment portfolio is defined as any collection of financial assets held by an investor). Twelve years later, Sharpe [Sha64], building on Markowitz's work, emphasised the importance of diversification in portfolio management, noting an explicit link between an asset's correlation to other assets in an investor's portfolio and that portfolio's overall risk and return potential. According to Sharpe, risk reduction through diversification increases if the portfolio's assets exhibit *a low degree of (or negative) correlation* to each other; conversely, should the level of correlation within a portfolio's constituents rise, its expected variance, and hence, the risk will rise even if there is no change to its composition [Sha64]. Despite its fundamental importance, the construction and management of well-diversified portfolios is highly challenging due to the great number of relations to analyse, which increases semi-quadratically with the number of securities held – an average-size investment portfolio (144 holdings) contains over 10,000 pairwise relations – and further complicated by the sheer volume of financial assets available for selection.

Adding to this complexity is the extremely volatile, dynamic nature of financial markets, particularly pronounced in times of market stress, when asset prices exhibit an elevated degree of co-movement (figure 2). Portfolio diversification may then break down just when needed most, further increasing the risk of losses to the investor [KHV*15, SS05, GSZ16]. In such volatile conditions, effective analysis of relationships within the assets becomes especially important but also challenging, as prices and behaviours change rapidly [GF00] and diversifying relations between assets become more valuable yet also scarcer. The task is complicated further by the presence of sub-structures whose behaviour over time varies and due to a variety of metadata (e.g. economic sector, position size within the portfolio) which is important to consider in the analysis.

The study of dynamically varying covariances and correlations is thus of great importance in many areas of the securities industry, including, *inter alia*, investment selection, portfolio construction, derivatives pricing, structuring and trading, financial index construction and the measurement and management of financial risk [SF12, GSZ16]. In current industry practice, however, correlations are often represented in a manner which requires considerable analytical effort from the user (see section 3.2). Moreover, the consideration of low correlations along with high correlations and across different subsets is often not supported. As a result, many risk managers and other users of correlations carry out a great deal of manual work, thus, making better decisions when working on such complex systems is only possible through the adoption of carefully designed, sophisticated techniques that facilitate the concurrent analysis of several dynamic relations.

In this paper, we propose a visual analytics approach to address challenges arising in current practice. Using the analysis of

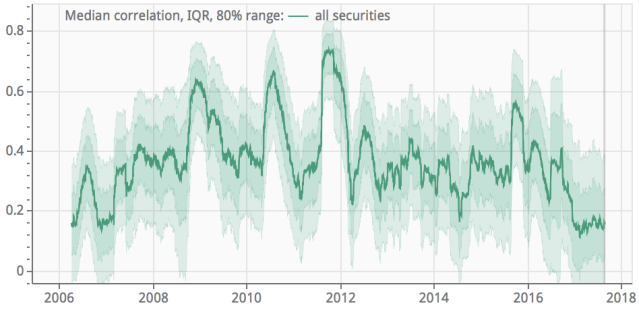

Figure 2: *Correlations between stock returns vary over time and are generally higher in times of market stress. Here, we see how the distribution of rolling three-month pairwise correlation coefficients (r) between S&P 500 (US large-capitalisation stocks) index constituents' daily price returns has fluctuated since 2006.*

dynamic correlations in securities markets as a motivating example, we design and build interactive visualisation solutions suitable for the simultaneous identification of both high and low correlation relationships (and, more generally, for the visual analysis of high-dimensional data with pairwise distances) and the tracking of changes therein over time. This provides an effective and robust solution that enables analysts to investigate the relations between correlations within investment portfolios and indices of up to several hundred securities for the exploration and analysis of changes in correlation structures over time.

Our solution is designed through a user-centred methodology by identifying the key requirements of the domain through expert interviews and an in-depth literature review. We adopt concepts such as multi-dimensional scaling, weighted complete graphs and threshold networks to present interactive, animated displays which use proximity as a visual metaphor for correlation and animation stability to encode correlation stability. We devise *context-aware interaction techniques* coupled with auxiliary views that enable analysts to interactively determine and investigate "sub-networks" of varyingly correlated securities. We identify and introduce *behaviour profiles* for securities that indicate the structures and behaviours of interests from a financial analysis perspective. These profiles are related to the analytical tasks and we revisit them in a number of analysis stories carried out as part of the validation of our approach. We further evaluate our method through measurement of the quality of layouts created by our algorithm and frame-rates of the resultant animations, as well as by a follow-up user-study involving interviews with experts. To summarise, the contributions of this work are:

- Observations and lessons learned from a design study to characterise the domain of financial correlation analysis
- Interactive visualisation techniques for the comparative analysis of sub-structures in dynamic networks
- A characterisation of key visual behaviours seen in dynamic weighted complete networks of financial data
- Use cases and follow-up interviews to demonstrate the effectiveness of interactive visual methods in the application domain

2. Related work

Investment portfolios often contain high numbers of holdings or securities, with US mutual funds typically holding between 40 and 120 stocks [SS05]. More recent work suggests that these numbers have increased, with a mean of 144 and a standard deviation of 271 stocks held [GSZ16]. The ability to visualise a substantial number of pairwise correlations concurrently is therefore of particular interest. This, however, results in data fidelity being a concern, with many methods discarding a large proportion of the correlation matrix data, making the identification of low-correlation pairs difficult. We review the representation techniques of correlation matrices under a number of categories in the following.

2.1. Hierarchical visualisation idioms

One popular approach is to postulate that the constituents of a portfolio, index or other group of stocks have a hierarchical structure driven by their pairwise correlation coefficients [Man99]. Once this structure has been determined, it is visualised using techniques such as dendograms, minimum spanning trees (discussed below) or Neighbour-Nets. A general criticism is that a hierarchical structure may be imposed where none exists [RR14]; furthermore, these approaches remove a great amount of detail from the visualisation of correlation matrices, which results in fairly simple visual representations but is problematic for our use case: an absent or distant hierarchical connection between a pair of securities is not a sufficient condition for the existence of a low correlation.

2.2. Proximity, similarity and distance metaphors

Correlations, as they measure the degree of linear dependency of a relationship, may be interpreted as measures of similarity [EL09] and hence of proximity between pairs of dimensions, their inverse thus a measure of distance. This family of visualisation techniques takes advantage of that concept. For instance, ordered heat maps of correlation matrices (i.e. distance matrices) are widely used and understood and often useful for the easy identification of clusters of correlated variables [MMGG16], but vary in effectiveness dependent on the quality of the ordering technique used [SS02]. Scatter plots based on multi-dimensional scaling (MDS) are used in many applications which analyse the similarity, relatedness or proximity of data [Man99, BGM12] – a simple visual representation of relationships between variables is created by projecting high-dimensional data into fewer (usually two or three) dimensions, whilst aiming to preserve pairwise distances. This idiom makes MDS suitable for visualising changes in correlations [GF00] and other network topologies [LS08, LSSdN08] over time. In a static setting, MDS has been used as a layout to communicate similarities between data dimensions as represented through their correlations [YPH*, TSL*17]. However, due to the stochastic nature of the computation, and its susceptibility to errors in data and oversensitivity to inherent structures [BGM12], MDS needs to be incorporated with care.

2.3. Pruned-graph techniques

Graphs (a.k.a. node-link diagrams, network diagrams) are a common visual metaphor for the depiction of correlations, as they are expressly intended for the illustration of relationships between entities [Kir16]. Graph-based approaches combine elements of both hierarchical and distance-based representations, and much of the research into graphs as visualisations of financial correlations has focused on the efficient reduction of graph size and complexity with filtering as the widely adopted approach [TLM10].

Examples of such methods, in descending order of severity of filtering, include minimum spanning trees (MST) [Man99,OCKK02], average linkage minimum spanning trees (ALMST) [TDMAM07], planar maximally filtered graphs (PMFG) [TADMM05], asset graphs (AG) [OCK*03a], and threshold networks (TN) [TLL10]. In the context of correlation analysis, a minimum spanning tree is a special case of an acyclic node-link diagram (a.k.a. network graph) where links only exist between closely-linked pairs of nodes [Man99]. It is defined as a simply-connected graph where all N nodes of the graph are connected by $N-1$ edges arranged such that the sum of all edge weights is minimized [OCKK02]. The acylicality requirement leads to some higher-correlation relationships between pairs of stocks being omitted whilst other, lower-correlation, pairs of stocks are maintained (in order that all stocks have at least one edge); the technique is known to favour strong positive correlations [BPS16]. MST's proponents argue that the technique creates meaningful hierarchical clusters when applied to various types of financial assets [OCK*03b, NRM07], but the method has been criticised for being *"probably the most severe form of data reduction"* [BPS16]. ALMST, PMFG, AG and TN address this criticism by extending MST's concept (and increasing the number of edges shown) in various ways – for example, TN have edges for all pairwise correlations above a user-defined threshold [TLL10] and AG show the n strongest correlations [OCK*03a]. In all pruned-graph methods, however, not all correlations between pairs of stocks are shown; the pruning algorithms cited use strength of correlation as inclusion criteria, not considering removed relations in layout calculation and therefore performing poorly at identifying low-correlation relationships.

2.4. Hybrids: weighted complete graphs/correlation maps

Correlation maps (CM) [ZMM12,ZMZM15], a.k.a. weighted complete graphs [QCX*07, PCB11], are a hybrid approach which attempt to address the deficiencies of the approaches reviewed above. Akin to MDS scatter-plots, CM present two-dimensional visualisations of variables where the distance between each pair of variables represents their similarity, taking advantage of the phenomenon that users naturally interpret closely located nodes in graphs as strongly related [DC98, QCX*07]. CM incorporate features from node-link graphs, being laid out using a mass-spring model where the forces determining edge length are driven by the pairwise Pearson correlation coefficient of the connected vertices. The method also draws on the concept of threshold networks: for greater clarity, links for correlations below an arbitrary threshold may be hidden in the visualisation, but are still taken into account in calculating the layout. CM are versatile and equally appropriate for high and low correlations; their topology not only shows relationships between pairs of variables, but also the overall relationship between all variables [QCX*07], aiming to preserve all information between pairs of variables in a correlation matrix.

2.5. Visualising changes in correlations over time

There exists a growing body of research in the field of dynamic networks [BBDW17]. Such visualisations tend to take one of four approaches: the presentation of aggregated temporal information in a single image; three-dimensional space-time cubes; series of images depicting evolution over time; or a step-by-step *"time multiplex"* display, sometimes with animated transitions between steps [BPF14, BDA*17, CRMH12]. The time-multiplex method offers several advantages, including the provision of cognitive support to users trying to understand the differences between steps, but also suffers from some limitations, including an increasing cognitive workload for the user as the number of changing elements increases. Time multiplexes also allow a greater level of detail to be achieved as only one (large) display is on screen at a given time, rather than partitioning the screen space into smaller plots (*"small multiples"*) for simultaneous display [AP16] or aggregating multiple periods into one static graphic.

3. Design process

Our work broadly follows Munzner's four-level nested model for visualisation design and validation [Mun09, Mun14]. We describe our work and findings for each layer of this model in the following sections, beginning by characterising the domain problem, with a view to formulating the analytical tasks to support, using a combination of literature review and primary research.

3.1. Domain problem characterisation: investigating the literature

As stated in section 1, there is broad agreement in the literature that periods of high market volatility, risk and stress are characterised by higher levels of correlations between securities than more "normal" market environments. Benefits of portfolio diversification are therefore reduced during high-volatility periods in financial markets, with this effect being particularly pronounced in falling markets [SS07]. As markets may be characterised as *volatile* some 20% of the time and *highly volatile* a further 10% of the time [GZ08], portfolio diversification failure may occur more frequently than market participants expect; given that one of portfolio diversification's aims is to reduce risks, there is a certain irony that it typically fails at those times when it is most needed. Researchers have raised the intriguing possibility that breakdowns in portfolio diversification might be anticipated by tracking the behaviour of stock correlations [PKS*12], using changing inter-market dependencies as early warnings of financial stress and crises [KRLBJ12].

In practice, correlations between stocks are often assumed to be constant [PKS*12]; however, numerous studies (e.g. [LS95, DD97, GF00, DGG*00, AC02, GLRG05, Pap14]) have demonstrated that correlations between securities' (and/or financial markets') returns are not static. Individual stocks' correlation behaviours can vary significantly between 'drawdown' and 'drawup' phases of markets; falling markets see greater correlations between stock returns, with rising markets exhibiting greater diversity [DGG*00]. This phenomenon is particularly pronounced in smaller companies, lower-beta (more defensive) and value stocks, and poorly-performing stocks [AC02]. Major market crashes often correspond with changes in markets' correlation structure [GF00]. Correlations are also a determinant of expected security returns and correlation risks are thus an important factor in securities pricing; defensive stocks which outperform *in times of high correlations* may be more attractive to investors, commanding higher valuations and thus lower expected returns [KPR09]. The structure of correlations between securities' returns is informative for returns of stocks and market indices and movements in currency rates; inter-market correlations were found to have both long-term ("slow") and short-term ("fast") dynamics, making analysis difficult, with the fast time-scale being about 60 trading days [STZM11]. Beyond all these, there are many other external factors that might influence the dynamics of the correlation relations, such as the nature of the companies' businesses, the extent to which their profitability is sensitive to the same economic variables, the countries in which they are listed, all of which make the analyses even more complex, requiring manual investigation by an expert.

Key findings: Critical tasks for an effective visual approach for the tracking of financial correlations include the ability to efficiently *gain an overview of correlation structure of a set of financial assets* (e.g. a market index or portfolio) and *identify changes therein*, such as whether correlations are rising or falling. Diversification efforts could be supported by the *analysis of pairs and larger groups of assets' dynamics*, with the effective *identification of low-correlation assets* as important as the *analysis of higher-correlated groups*.

3.2. Domain problem characterisation: interviews with practitioners

Our next step was to establish more detailed requirements from users and practitioners of correlation analyses in the securities industry. We therefore conducted five interviews with London-based users of correlation analyses (coded **P1**...**P5**): two risk managers, two quantitative analysts and the manager of a quantitative investment fund. In the course of one-hour meetings, participants answered three Likert-scaled questions and eleven open-ended questions. The scaled questions were designed to validate the importance of the topic and measure the perceived need for improvement; the open-ended questions were formulated to further validate topic importance, understand the use cases, determine current practice and specific areas for improvement, and establish the most important software features required. Interviews were recorded and a thematic analysis was undertaken.

We found that correlation analyses in investment management and financial markets are seen as very important by users, with the ability to track changes over time at least as important as the understanding of absolute levels of correlation. All participants expressed some perceived need for improvement to their current approach to this task. A great variety of requirements and needs were mentioned in the interviews; as a result, flexibility of use cases is an important consideration. Financial correlation analyses are carried out for portfolio construction, risk budgeting, risk management/analysis and investment research purposes; further use cases include factor and scenario analyses, the discovery of portfolio skews and ex-post analyses to explain drawdowns. Interviewees would like to be better able to *identify clusters of highly-correlated stocks*, but also to *find uncorrelated assets*. Analytical results are produced for dif-

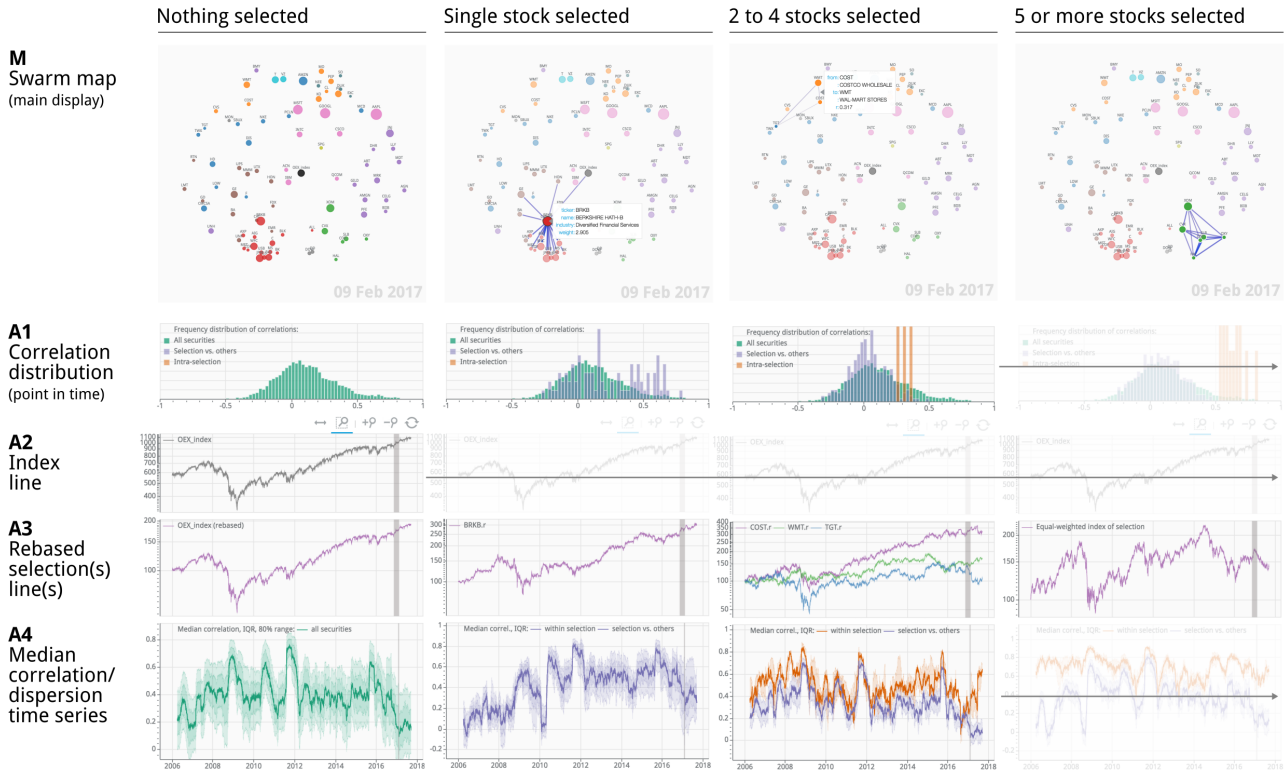

Figure 3: *Context-sensitive interactive displays in the c\swarm user interface. Main and auxiliary displays change depending on whether, and how many, nodes are selected. For example, in the main display, selecting only one node switches on all links between that stock and others (subject to interactive filtering). Selecting two or more nodes shows only links between the highlighted stocks. Interactive tooltips show further detail when hovering over nodes and edges. The analysis time window is moved by double-clicking on any of the line displays.*

ferent audiences with varying levels of analytical sophistication; simplicity, ease of interpretation, ease of use and user interface responsiveness (performance) were widely seen to be desirable. The ability to *link in further attributes of financial assets* (e.g. position size, market value, industry sector) was also seen as important.

Currently-used analytical approaches were seen to suffer from several deficiencies; the complexity of current representations was a recurring theme. Users noted issues such as *the instability of financial correlations* over time and in different market environments and the *difficulty of identifying securities with low and/or stable correlations*. Popular techniques appear particularly ill-equipped to deal with periods of high stress in markets, when the *ability to track correlations* is arguably most important. There also does not appear to be an accepted 'standard' way of visualising financial correlations beyond the tabular correlation matrix for point-in-time views, sometimes represented as a heat map, and line plots of pairwise correlations over time. Whilst the latter are widely used and understood, respondents agreed that the large number of plots needed, given the number of positions in a typical investment portfolio, mean that this is not practicable for monitoring shifts in correlation across an entire portfolio. Factor analysis, where securities are not correlated with each other, but with thematic factors, was mentioned as a useful method, but seen to only be tractable if few factors are used. As a result of these deficiencies, a great amount of manual work is carried out to track changing asset correlations.

3.3. Operation abstraction: analytical tasks

A synthesis of our findings from the literature review and primary research allows the specification of four analytical tasks (**T1**...**T4**) to be supported by our approach:

- **T1.** provision of a graphical overview of the correlation structure of a portfolio, index or other group of stocks or other financial instruments at a given point in time
- **T2.** the analysis of changes in the correlation structure over time and the nature and direction of such changes
- **T3.** the investigation of highly-correlated pairs and groups (clusters) of assets which contribute to portfolio risk; and
- **T4.** the investigation of uncorrelated or low-correlated securities to support portfolio risk reduction by diversification

3.4. Visual encoding and interaction techniques

Our approach's visual metaphor, the *proximity swarm*, combines concepts from weighted complete graphs (correlation maps), multi-

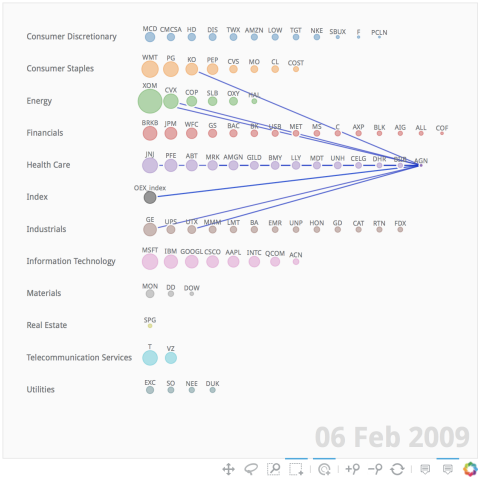

Figure 4: *The **'Sorter'** view rearranges the main display by sector and size to support easy selection of nodes and groups; the full range of interactions and animation is available in this mode.*

dimensional scaling and threshold networks in order to address the concerns identified with current methods. It is named for its use of proximity to encode correlations and its resemblance to a swarm of flying insects when animated.

We use node-link graphs as the underlying visual idiom for the main display. Each datum (node) represents one security or financial instrument (such as a stock, index, portfolio position, bond, currency, etc); its market value is encoded as the glyph's area and its sector by glyph colour. To address the deficiencies of pruned-graph approaches, the display is laid out as a complete graph (i.e. all nodes are connected to all other nodes). Edge weights represent (the strength of) pairwise correlations at a given, common point in time; as all matrix dimensions have a correlation value versus all other matrix dimensions at all times, it is the edge weights that change over time, not their presence or absence (unlike in conventional dynamic network-based idioms). This makes the display of links between nodes (edge glyphs) redundant, so they are not drawn by default; their presence or absence becomes an available channel for encoding other information (for instance, the pairwise correlation being over a certain threshold [TLL10, ZMM12, ZMZM15]). We encode correlations between financial instruments by the distances between node glyphs, with it being desirable that two arbitrarily chosen stocks which are highly correlated are closer to each other than two other stocks which are less highly correlated. This follows the weighted complete graph/correlation map visual idiom [QCX*07, ZMM12, ZMZM15] and preserves the maximum amount of information from the correlation matrix. The visual idiom is selected for its conceptual simplicity, intuitive nature and likely ease of interpretation by non-technical audiences. It is intuitive given that *"proximity implies relationship"* [GFV13] and follows the UI design trope that *"similar things should look similar, different things different"* [SM86]; users naturally interpret closer nodes on a graph as being more strongly related [DC98].

The passage of time is represented in the proximity swarm using a time-multiplex method to allow room for a sufficiently large display to visualise large numbers of assets, with auxiliary linked displays presenting the context. A visual encoding which follows naturally from this approach is the extent to which nodes move. During periods where correlations are stable, the display is stable; conversely, in times of rapidly shifting correlations, the display exhibits a great deal of movement.

Additional details on demand [Shn96] are shown in interactively summoned visual encodings in the main display **M** and four *context-sensitive auxiliary plots* **A1-A4** (figure 3), which also provide evidence that the behaviours described in the next section are not merely visual artefacts of the stochastic layout technique used. When one node is selected, links between that node and all others are shown; when multiple nodes are selected, all links between the selected nodes are shown, but no others. Links may be filtered interactively and are triple-encoded for strength of correlation using a diverging blue (strongly positive)–grey (zero)–red (strongly negative) colour scale as well as line width and transparency; visual 'pop-out' of highlighted nodes is provided by increasing the transparency of non-selected nodes. To further reduce visual clutter, node labels and legend may be hidden by the user; easy selection of nodes is supported by an additional display mode, the *'Sorter'*, which re-arranges the nodes by sector and size (fig. 4). The four auxiliary views display frequency distribution(s) of correlations at the point-in-time analysis window (**A1**), price time-series for the index/portfolio analysed (**A2**), price time-series of any items selected (**A3**) and time-series of central tendency and dispersion of correlations (**A4**). Elements shown in these views change depending on whether and how many nodes are selected (fig. 3).

3.5. Layout algorithm design, computational framework and implementation: the c|swarm tool

As discussed above, we use the strength of correlation to directly drive the layout of our visualisation rather than merely as binary criteria for the existence or absence of links between nodes. In our prototype, pairwise Pearson correlation coefficients between stock returns are pre-calculated from time-series of percentage changes in daily prices upon starting the software; a rolling, constant 65-weekday time window (i.e. three-month rolling daily correlations) is used to capture "fast" correlation dynamics [STZM11] and limit the time taken to initially calculate the source data cube of correlation matrices versus time. Any other parameter set or temporally-varying pairwise distance measure could be adopted here. For layout calculations, where R is a matrix of pairwise correlation coefficients for any given time window, distance matrices are calculated simply as $1-R$, with negative correlation coefficients therefore shown as greater distances than positive correlation coefficients: in our use case, the identification of strongly negatively-correlated pairs of stocks is highly desirable. If the strength of the relationship is more important than its direction and/or direction is encoded using another visual variable, the distance matrices may be calculated as $1-|R|$ or $1-R^2$.

The layout algorithm translates this distance matrix into a visual representation. We identify the layout algorithm for the swarm using two experiments. In the first experiment, we evaluate seven different candidate algorithms for calculation time, layout quality and

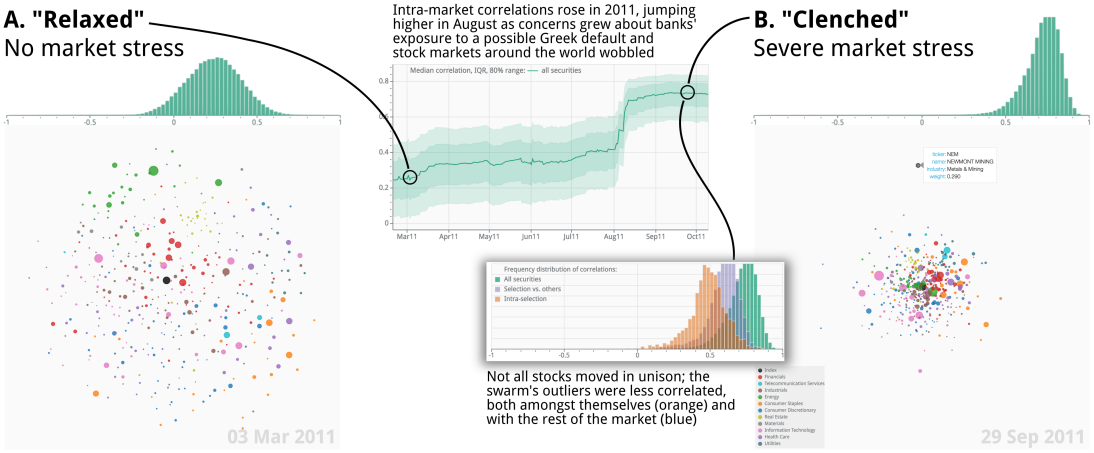

Figure 5: *The swarm "clenches" (**B1**) in times of market stress (such as the 2011 debt default crisis, **right**): when correlations are high, less-correlated stocks are of particular interest. Many of the stocks in the periphery are in 'defensive' sectors such as Consumer Staples, Health Care, Technology and Utilities; Newmont Mining, a 'safe haven' gold mining stock, is a great distance from the rest of the swarm.*

stability of layout quality over time using a small test data set (31 series) that span an 18-month period (see Appendix, Sec.6 for details). Layout quality is measured using Pearson and Spearman correlation coefficients between vectorized correlation matrices and arrays of pairwise Euclidean distances between the nodes in the calculated layouts – conceptually similar to the *Stress-1* measure used conventionally in MDS [BGM12]. A second round of parameter tweaking tested different configurations of the three best algorithms from the first experiment over five sets of test data covering 3,000 trading days between January 2006 and September 2017.

The concept of force direction, where distances between nodes are informed by forces specified by edge weights, would seem highly relevant, but classical force-directed methods [Ead84,FR91] often also aim to adhere to aesthetic rules such as minimising edge crossings, layout symmetry, uniform edge lengths and uniform node distribution [GFV13] which are not appropriate here. By contrast, dimension-reduction techniques which project pairwise distances between items in high-dimensional space to lower-dimensional space (e.g. the two dimensions of a flat plot) are useful for laying out complete graphs, and in particular, Multi-Dimensional Scaling (MDS) techniques [BGM12] are well suited to our task given their explicit aim of preserving distances. A recursive implementation of metric MDS, where each time slice is calculated using the previous time slice's layout as a seed, emerged as most suitable for the task across all performance criteria. This approach substantially improves calculation times, allowing online calculation of time-multiplex animations at usable frame rates for test data sets numbering over 400 stocks. Layouts are cached once calculated, ensuring that each point-in-time correlation matrix is assigned one layout for the full duration of a session of using the software, mitigating the effects of the stochastic nature of MDS computations. Further incremental performance gains are achieved by vectorizing the distance matrices, taking advantage of distance matrices' inherent symmetry for efficient storing and processing. There exists the potential to further enhance performance and scalability by applying progressive analysis techniques [FP16,TKBH17] to the layout algorithm design, for instance by starting the calculation of the layout cache immediately upon opening the software and/or by carrying out calculations in auxiliary displays **A2** and **A3** at progressively finer levels of detail when nodes are selected.

3.6. Mapping behaviour profiles

Using the distance between nodes as the main encoding idiom for correlations (and driving the layout with this encoding) in combination with a time-multiplex idiom results in a number of behaviours of interest that also have a semantic mapping from a financial analysis perspective. Here, we identify and introduce these as a series of *behaviour profiles* (numbered **B1**, **B2**, ...) to support the analytical tasks **T1**-**T4** as identified in section 3.3; examples of these profiles in action are described in section 4.2.

In times of market stress, when general correlation levels rise and dispersion falls (**T1**, **T2**), the entire swarm *clenches* tightly together (**B1**, fig. 5), supporting the easy identification of less-correlated stocks as outliers. As market stress eases and correlations fall, the swarm *relaxes* (**B2**, fig. 5), with distances between nodes increasing. Highly inter-correlated groups of stocks are shown as tightly clustered (*bunched*; **B3**, fig. 6) groups of nodes; assets and/or groups which are relatively highly correlated to other stocks can be found *closer* to the swarm's centre. *Central clustering* (**B4**, upper panel, fig. 8) occurs when a group of stocks is highly correlated both within itself and to the rest of the market; the group's nodes are clustered in a central location, where distances to other nodes are shorter. Conversely, nodes found on *opposing* sides of the display likely have low or strongly negative correlations with each other (**T3**). Stocks or groups of stocks which are generally less correlated to others (**T4**) are situated near the edge of the swarm display, *further away* from most other nodes, their exact location determined by which stocks their returns most closely correlate with. *Peripheral bunching* (**B5**, fig. 1), where a tightly clustered group of stocks

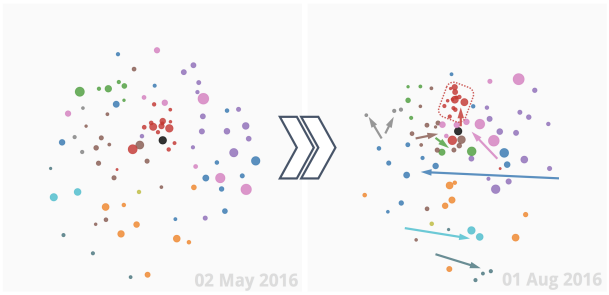

Figure 6: ***Rotation*** *(**B9**): from May to August 2016, we see a great deal of movement in nodes' positions relative to each other, suggesting that the underlying correlation structure has changed substantially. Many financials stocks are closely* ***bunched*** *(**B3**; red dotted line), indicating strong correlations within that group.*

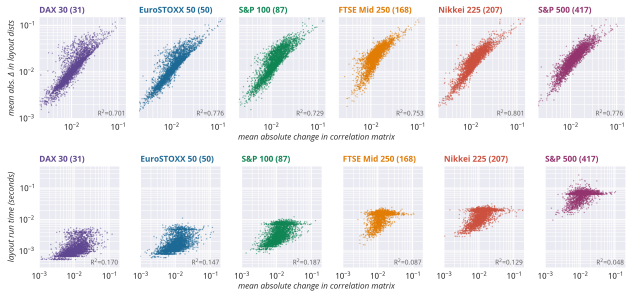

Figure 7: ***Top:*** *The layout algorithm effectively encodes correlation stability as layout stability in all test data. Number of nodes in parentheses.* ***Bottom:*** *Layout calculation speed is driven by the number of nodes in the data with a weak response to change in the correlations within the frames.* ***Note:*** *Log scales used on all axes.*

is found near the edge of the swarm, occurs when a (probably homogeneous) group of stocks is highly correlated within itself but less correlated to the rest of the market; by contrast, if a sector's nodes are *broadly dispersed* (**B6**, lower panel, fig. 8) across the swarm, the sector's correlation structure is likely to be heterogeneous and similar to that of the market as a whole. When investigating changing correlation structures over time, stocks with low and changing correlations to the rest of the swarm *orbit* (**B7**, fig. 9) around the outside of the display; stocks which are more highly correlated to the rest of the swarm, but whose correlations are changing, *meander* (**B8**, fig. 9) through the swarm. When there is a discontinuity in stock behaviour within a market leading to changes in correlation structure over a short period (*rotation*; **B9**, fig. 6), node positions change rapidly, with the swarm animation sometimes appearing to 'jerk'; during more stable times, most nodes move only gradually. Layout stability is thus used as a visual encoding for correlation stability.

4. Validation and evaluation

4.1. Scalability and layout quality evaluation

We tested the quality of layouts calculated by our algorithm by capturing statistics for full layout time-multiplex runs over eleven and a half years of daily data for the five test data sets as well as one further set of previously unseen data. The algorithm produces high-quality layouts, with stability over time of the display correlating well to the stability (or otherwise) of the test correlation data. We found that the layout algorithm produced particularly high-quality layouts in times of high market volatility and stress; the encoding of correlation stability as layout stability was also visualised with a high degree of fidelity (upper panel, fig. 7). Layout *stability* proved to have a statistically insignificant link to levels of market volatility; we observed that during times of market stress, the animation (and the underlying correlation structure) was often remarkably stable. Average animation frame rates ranged from 71.0fps for our smallest test dataset (with 31 nodes) to 13.4fps for the largest (417 nodes), with layout calculation times being primarily driven by the number of nodes rendered and to a lesser extent by the stability of the underlying correlation matrices (lower panel, fig. 7). Using pre-calculated layouts (to gauge the underlying rendering engine's performance), the largest test dataset was animated at an average frame rate of 65.7fps. Extrapolating from the performance on the largest test data set, the frame rates reported should be achievable for approximately 85% of investment portfolios, assuming that portfolio sizes are normally distributed and that estimates from Goldman et al. [GSZ16] generalise.

4.2. Validation of visual encoding: analysis stories

Qualitative tests of the tool's effectiveness were conducted by visually inspecting and interacting with the full animations for each of the six layout validation data sets and intermittently randomly selecting nodes or groups of nodes and inspecting the correlation coefficients corresponding to the highlighted edges. These tests formed a basis for validation of the visual encoding and investigating its appropriateness for visual storytelling. To validate the visualisation's ability to support the analytical tasks and demonstrate occurrences of its expected behaviours, several case studies (analysis stories) were identified to address suitability for the analytical tasks specified in the design stage. (Refer to Appendix section 5 for further example analysis stories).

Analysis Story 1: In fig. 5, we see an example of *clenching* and *relaxing*. In March 2011, US large-cap stocks' median correlation was low; nodes are spread out across the display and the graph is *relaxed* (**B2**). By late September, markets around the world wobbled as it seemed likely that Greece would default on its sovereign debt, with knock-on ramifications for the wider Eurozone and global economies [Wik18]. Correlations increased to the highest level seen in any of our test data sets and remained elevated for some time afterwards; the swarm plot shows a tightly *clenched* (**B1**) group of stocks. Nodes from the periphery of the swarm disproportionately represented 'defensive' sectors such as consumer staples, utilities, health care and technology; selecting them shows us that these stocks had distinct correlation characteristics from the rest of the index at that time (multiple histogram, inset, fig. 5).

Analysis Story 2: In figs. 1 and 8, we present three different US market sectors to illustrate *peripheral bunching* (**B5**), *central clustering* (**B4**) and *broad dispersion* (**B6**). The utilities sector (electric-

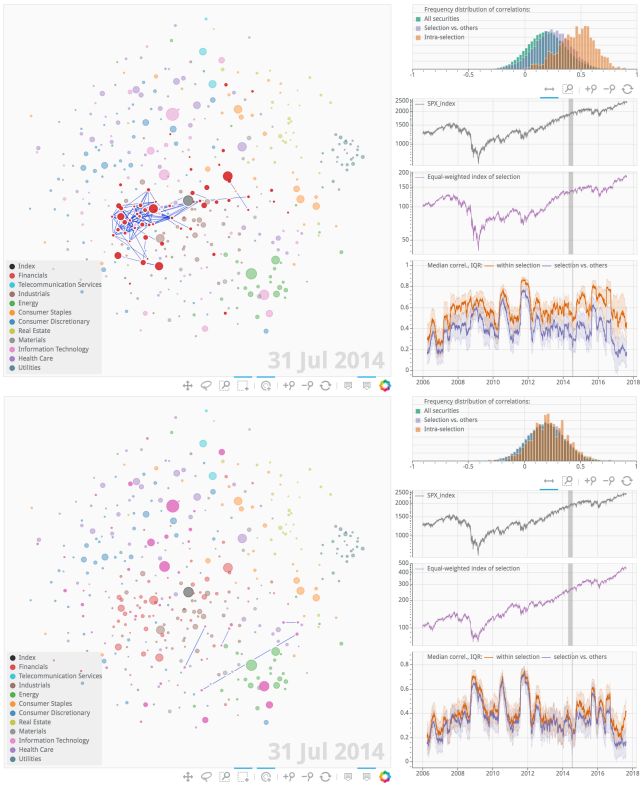

Figure 8: *Industry sectors with different levels of homogeneity exhibit different behaviours in the swarm (links for $r < 0.7$ filtered out in all examples).* ***Upper:*** *Central clustering (**B4**): stocks in the financials group (red) are less tightly bunched than utilities or energy stocks, but occupy a central location in the swarm.* ***Lower:*** *Broad dispersion (**B6**): IT sector (magenta) stocks are scattered throughout a wide area of the swarm. The histogram shows that their correlation structure is similar to that of the entire market's.* ***In fig. 1:*** *Peripheral bunching (**B5**): the utilities sector exhibits high intra-sector correlation and low correlation to other stocks.*

ity and gas distributors) is quite homogeneous, as its constituents' performance is subject to very similar factors: government regulation, input commodity costs and interest rates (as the business is capital intensive). However, the sector is less dependent on general economic health than others, so its correlation to others is low. Visually, this means that the utilities are located on the edge of the swarm, tightly clustered together and at a distance from other stocks (**B5**). The financials sector exhibits *central clustering* (**B4**): the nodes are less tightly spaced, as the sector is more heterogeneous, containing a number of different industries (retail banks, insurers, other financial services). The stocks' fairly central location within the swarm suggests that they are more correlated to stocks from other sectors, consistent with the link between the financial sector's fortunes and those of the wider economy (both directly and via interest rate sensitivity). Finally, the technology sector contains a diverse range of stocks which, at the analysis window shown, were not highly intercorrelated and thus *broadly dispersed* (**B6**) across the swarm; this suggests that at the time, stock-specific factors were more important than sector-level themes in determining behaviour, and returns were dispersed.

Analysis Story 3: We observe examples of *orbiting* (**B7**) and *meandering* (**B8**) in fig. 9, which follows Apple's movement through the S&P 100 index over the course of a few months. In a time where market correlations were generally stable and the swarm was *relaxed*, Apple's position around the edge of the swarm tells us that the stock exhibits low correlation to most others; the node's movement over time illustrates that its relationship with the rest of the market keeps changing. After a year of *orbiting*, the stock's behaviour changes, and the node *meanders* through the swarm as its correlation to the rest of the market rises; by mid-2015, Apple's node is located near the middle of the swarm (close to the market index, which it was the largest component of by then).

4.3. User evaluation

To further evaluate our approach's suitability to support the tasks from section 3.3, we re-visited the participants from our requirements-gathering study (section 3.2) and carried out a second series of interviews. In the first part of each interview, following a brief software demonstration, participants were set five standardised exercises **E1**-**E5** covering analytical tasks **T1**-**T4** outlined in section 3.3: to identify and analyse a highly correlated group of stocks (**E1: T1, T2, T3**), to compare two highly correlated groups of stocks (**E2: T2, T3**), to find a period of unusually high correlations and identify outliers in that period (**E3: T1, T2, T4**), and to find and analyse pairs of low-correlation (**E4: T2, T4**) and high-correlation (**E5: T2, T3**) stocks. Each exercise was allocated four minutes, with tasks not completed in time marked incomplete. Participants were asked to "think aloud" whilst performing the tasks, with a particular focus on any insights reached, and task performance observed and recorded by the interviewer. After the structured tasks, participants were given the opportunity to further explore the visualisation, in order to gain further observational, anecdotal evidence. To conclude, participants answered a list of seven open-ended questions to cross-check observations from the timed tasks, understand strengths and weaknesses of our approach and identify areas for further development. Details and results of the study are in Appendix 1, section 3.

Performance by research participants was generally strong, with most exercises being completed in under two minutes. Cluster identification exercises (**E1**, **E2**) took more time despite participants reporting that they found them easily, with overview/outlier tasks (**E3**) problematic for one participant. There were several examples of sub-tasks being answered as soon as the researcher set the task, with most participants completing all exercises in substantially less than the allocated time of four minutes, and only one occurrence of a task not being completed by a participant in the allocated time. Size and colour encoding and the highlighting of selections through the presence of links has shown to be effective, although one user mentioned that finer control of selections would be desirable. Without exception, the study participants were positive about the visualisation approach and mentioned its potential to be highly useful. Strongly positive feedback from the study participants suggested that most gained insights would have been difficult to achieve using the diverse range of tools used in current approaches, although one

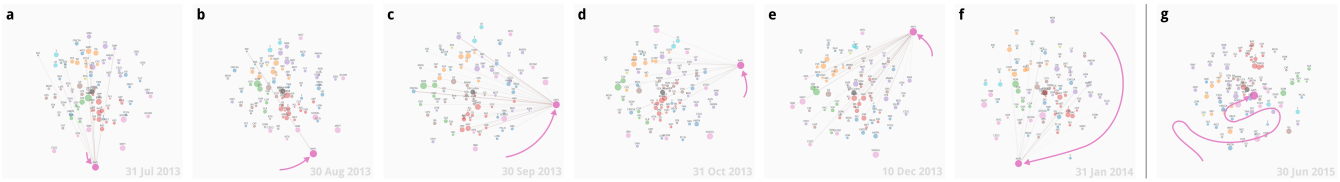

Figure 9: *"Orbiting:" in a period where the swarm was generally quite stable (a-f), Apple (AAPL; magenta; edges for positive correlations filtered out) "orbited" (**B7**) around the plot, indicating low, variable correlations to other stocks. By mid-2015 (g), the correlation behaviour had changed and the stock had "meandered" (**B8**) to a position near the swarm's centre.*

participant also stated that they found the animation "hard to get to grips with". In particular, the illustration of the extent to which correlations converge in financial crises [BL09] was greeted with surprise and interest by all research participants, even though all but one had been working in investment management throughout the 2008 and 2011 crises (thus experiencing them first-hand); this suggests that our approach also shows promise for visual storytelling and education. Study participants described the approach as *"innovative and novel"* (**P1**), *"so much better than looking at correlation matrices"* (**P3**), and told us that they *"cannot think of a way of capturing the dynamics of correlations as powerfully as [we] have"* (**P2**). In one instance, a participant was unable to finish an exercise in the four minutes allocated. This may partly be attributable to the interviewer's lack of experience, and also to a preference on a less visual and more numerical representation. This is a sign that potential users need to be given sufficient time and training to familiarise with these concepts.

5. Discussion

The provision of static elements to support the animation's effectiveness proved helpful to users, echoing findings that hybrid approaches using both animation and time lines may prove useful for task performance in certain contexts and that users may find them helpful [RM13]. The utility of good interaction design in helping users to deal with perceptual difficulties from overly fast or complex animations [TMB02] was demonstrated. One critique here could be that the animated representation could easily lead to the perception of behaviours that are not necessarily grounded but are solely artefacts of the layout algorithm. This is a valid concern – in order to mitigate such issues, we conducted rigorous tests during layout algorithm identification and validation, and provide context-aware auxiliary views with controlled animation features to help support the verification of any behaviour identified.

Existing methods of visualising financial correlations may have too rigid an emphasis on filtering heuristics for the information contained in a correlation matrix for the sake of identifying a hierarchy. This makes them broadly unsuitable for risk management applications, as they cannot be used to find low-correlation assets. In the field of dynamic networks, much of the corpus of work operates on graph-theoretic measures such as node degrees or centralities in quantifying and analysing changes in graphs. However, in applications where strengths of relations are constantly changing, such approaches can have weaknesses that affect the interpretation of the metaphor – in particular the potential over-reliance on the information contained in the existence or absence of a link between two nodes. Our approach recognizes this through animated weighted edges, but further research on the analysis of the dynamics of weighted complete graphs with an emphasis on changes in their weights and attributes is required.

One key process in our design phase is the identification and utilisation of *behaviour profiles*. These are loosely defined, but semantically relevant, characteristics patterns that have the potential to be instrumental in future analyses conducted in our approach, for instance, in providing visual guidance to identify areas/periods of interest or externalising and interpreting patterns. Further work is needed to formally, i.e. through formulaic settings, define these profiles to enable their semi-automated extraction. Despite this shortcoming, we observe that the behaviour profiles serve as effective heuristics to conceptually guide and structure financial correlation analysis executed in our approach; we see potential merit in investigating their generalisability over a wider range of domains.

Currently, our approach supports the visual investigation and comparison of a single subset selection. One promising potential future work that emerged during the interviews is to allow the comparison of different groups of assets. Our current design of auxiliary views and interactions is not optimised to handle more groups and a further design exercise is needed to address that; we deem this to be a promising next step. The deployment of progressive analysis techniques [TKBH17] to improve the tool's scalability is another interesting area for further work.

6. Conclusion

This paper introduces a visual analytics framework for the interactive analysis of relations and structures in dynamic, high-dimensional correlation data, in particular in the context of financial data analysis. The adopted visual metaphor and interaction design are rooted in the literature and supported by research on current industry and academic requirements. We have demonstrated the technique's potential by validating the layouts created by the algorithm, illustrating interesting analysis cases and carrying out initial user evaluations. We introduced a series of behaviour profiles to help structure the analytical process and further supported analysis with context-aware views and enhanced interaction schemes. With our focus on information fidelity and distance preservation in the representation of financial correlations, we believe that our tool will be useful in the study of changing correlation structures and of periods of market stress, when correlations are high and identifying diversifying assets is especially important.

Acknowledgements

We wish to thank the research study participants for their time, warm reception and candid and valuable insights; our thanks also go to the numerous colleagues and friends who provided valuable feedback on our early prototypes.

This work would not have been possible without the work of the open-source software community (particularly the developers of and contributors to *Python*, *Bokeh*, *pandas*, *NumPy*, *SciPy*, and *scikit-learn*), whom we thank for their continuing dedication to the cause of developing great technologies freely available to all.

References

[AC02] ANG A., CHEN J.: Asymmetric correlations of equity portfolios. *Journal of Financial Economics 63*, 3 (2002), 443. 4

[AP16] ARCHAMBAULT D., PURCHASE H. C.: Can animation support the visualisation of dynamic graphs? *Information Sciences 330* (2016), 495–509. 4

[BBDW17] BECK F., BURCH M., DIEHL S., WEISKOPF D.: A taxonomy and survey of dynamic graph visualization. *Computer Graphics Forum 36*, 1 (2017), 133. 4

[BDA*17] BACH B., DRAGICEVIC P., ARCHAMBAULT D., HURTER C., CARPENDALE S.: A descriptive framework for temporal data visualizations based on generalized space-time cubes. In *Computer Graphics Forum* (2017), vol. 36, Wiley Online Library, pp. 36–61. 4

[BGM12] BORG I., GROENEN P. J., MAIR P.: *Applied Multidimensional Scaling*. Springer Science & Business Media, 2012. 3, 7

[BL09] BAUR D. G., LUCEY B. M.: Flights and contagion—an empirical analysis of stock-bond correlations. *Journal of Financial Stability 5*, 4 (2009), 339. 10

[BPF14] BACH B., PIETRIGA E., FEKETE J. D.: GraphDiaries: Animated transitions and temporal navigation for dynamic networks. *IEEE Transactions on Visualization and Computer Graphics 20*, 5 (2014), 740. 4

[BPS16] BIRCH J., PANTELOUS A. A., SORAMÄKI K.: Analysis of correlation based networks representing DAX 30 stock price returns. *Computational Economics 47*, 4 (2016), 501. 3

[CRMH12] CHUANG J., RAMAGE D., MANNING C., HEER J.: Interpretation and trust: Designing model-driven visualizations for text analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2012), ACM, pp. 443–452. 4

[DC98] DENGLER E., COWAN W.: Human perception of laid-out graphs. In *International Symposium on Graph Drawing* (1998), Springer, pp. 441–443. 3, 6

[DD97] DARBAR S. M., DEB P.: Co-movements in international equity markets. *Journal of Financial Research 20*, 3 (1997), 305. 4

[DGG*00] DROZDZ S., GRÜMMER F., GORSKI A. Z., RUF F., SPETH J.: Dynamics of competition between collectivity and noise in the stock market. *Physica A: Statistical Mechanics and its Applications 287*, 3 (2000), 440. 4

[Ead84] EADES P.: A heuristic for graph drawing. *Congressus Numerantium 42* (1984), 149–160. 7

[EL09] EGGHE L., LEYDESDORFF L.: The relation between Pearson's correlation coefficient r and Salton's cosine measure. *Journal of the Association for Information Science and Technology 60*, 5 (2009), 1027–1036. 3

[FP16] FEKETE J. D., PRIMET R.: Progressive Analytics: a computation paradigm for exploratory data analysis. *ArXiv e-prints* (2016). 7

[FR91] FRUCHTERMAN T. M. J., REINGOLD E. M.: Graph drawing by force-directed placement. *Software: Practice and Experience 21*, 11 (1991), 1129. 7

[GF00] GROENEN P. J. F., FRANSES P. H.: Visualizing time-varying correlations across stock markets. *Journal of Empirical Finance 7*, 2 (2000), 155. 2, 3, 4

[GFV13] GIBSON H., FAITH J., VICKERS P.: A survey of two-dimensional graph layout techniques for information visualisation. *Information Visualization 12*, 3-4 (2013), 324–357. 6, 7

[GLRG05] GOETZMANN W., LI L., ROUWENHORST K., GEERT: Long-term global market correlations. *The Journal of Business 78*, 1 (2005), 1–38. 4

[GSZ16] GOLDMAN E., SUN Z., ZHOU X.: The effect of management design on the portfolio concentration and performance of mutual funds. *Financial Analysts Journal 72*, 4 (2016), 49. 2, 3, 8

[GZ08] GEYER A., ZIEMBA W. T.: The Innovest Austrian pension fund financial planning model InnoALM. *Operations Research 56*, 4 (2008), 797–810. 4

[KHV*15] KENETT D. Y., HUANG X., VODENSKA I., HAVLIN S., STANLEY H. E.: Partial correlation analysis: applications for financial markets. *Quantitative Finance 15*, 4 (2015), 569. 2

[Kir16] KIRK A.: *Data Visualisation: a handbook for data driven design*. Sage, London, 2016. 3

[KPR09] KRISHNAN C. N. V., PETKOVA R., RITCHKEN P.: Correlation risk. *Journal of Empirical Finance 16*, 3 (2009), 353. 4

[KRLBJ12] KENETT D. Y., RADDANT M., LUX T., BEN-JACOB E.: Evolvement of uniformity and volatility in the stressed global financial village. *PloS one 7*, 2 (2012), e31144. 4

[LS95] LONGIN F., SOLNIK B.: Is the correlation in international equity returns constant: 1960-1990? *Journal of International Money and Finance 14*, 1 (1995), 3. 4

[LS08] LEYDESDORFF L., SCHANK T.: Dynamic animations of journal maps: indicators of structural changes and interdisciplinary developments. *Journal of the American Society for Information Science and Technology 59*, 11 (2008), 1810. 3

[LSSdN08] LEYDESDORFF L., SCHANK T., SCHARNHORST A., DE NOOY W.: Animating the development of social networks over time using a dynamic extension of multidimensional scaling. *El profesional de la información 17*, 6 (2008). 3

[Man99] MANTEGNA R. N.: Hierarchical structure in financial markets. *The European Physical Journal B - Condensed Matter and Complex Systems 11*, 1 (1999), 193. 3

[Mar52] MARKOWITZ H.: Portfolio selection. *The Journal of Finance 7*, 1 (1952), 77. 2

[MMGG16] MCKENNA S., MEYER M., GREGG C., GERBER S.: s-CorrPlot: an interactive scatterplot for exploring correlation. *Journal of Computational and Graphical Statistics 25*, 2 (2016), 445. 3

[Mun09] MUNZNER T.: A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 921. 4

[Mun14] MUNZNER T.: *Visualization Analysis and Design*. CRC Press, Boca Raton, 2014. 4

[NRM07] NAYLOR M. J., ROSE L. C., MOYLE B. J.: Topology of foreign exchange markets using hierarchical structure methods. *Physica A: Statistical Mechanics and its Applications 382*, 1 (2007), 199. 3

[OCK*03a] ONNELA J.-P., CHAKRABORTI A., KASKI K., KERTESZ J., KANTO A.: Asset trees and asset graphs in financial markets. *Physica Scripta 2003*, T106 (2003), 48. 3

[OCK*03b] ONNELA J. P., CHAKRABORTI A., KASKI K., KERTESZ J., KANTO A.: Dynamics of market correlations: taxonomy and portfolio analysis. *Physical Review E 68*, 5 (2003), 056110. 3

[OCKK02] ONNELA J. P., CHAKRABORTI A., KASKI K., KERTIÃL'SZ J.: Dynamic asset trees and portfolio analysis. *The European Physical Journal B - Condensed Matter and Complex Systems 30*, 3 (2002), 285. 3

[Pap14] PAPAVASSILIOU V. G.: Cross-asset contagion in times of stress. *Journal of Economics and Business 76* (2014), 133. 4

[PCB11] PICCARDI C., CALATRONI L., BERTONI F.: Clustering financial time series by network community analysis. *International Journal of Modern Physics C 22*, 01 (2011), 35–50. 3

[PKS*12] PREIS T., KENETT D. Y., STANLEY H. E., HELBING D., BEN-JACOB E.: Quantifying the behavior of stock correlations under market stress. *Scientific Reports 2* (2012), 752. 4

[QCX*07] QU H., CHAN W.-Y., XU A., CHUNG K.-L., LAU K.-H., GUO P.: Visual analysis of the air pollution problem in Hong Kong. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007). 3, 6

[RM13] RUFIANGE S., MCGUFFIN M. J.: DiffAni: Visualizing dynamic graphs with a hybrid of difference maps and animation. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2556. 10

[RR14] REA A., REA W.: Visualization of a stock market correlation matrix. *Physica A: Statistical Mechanics and its Applications 400* (2014), 109. 3

[SF12] SANDOVAL L., FRANCA I. D. P.: Correlation of financial markets in times of crisis. *Physica A: Statistical Mechanics and its Applications 391*, 1 (2012), 187. 2

[Sha64] SHARPE W. F.: Capital asset prices: a theory of market equilibrium under conditions of risk. *The Journal of Finance 19*, 3 (1964), 425. 2

[Shn96] SHNEIDERMAN B.: The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings 1996 IEEE Symposium on Visual Languages* (1996), p. 336. 6

[SM86] SMITH S. L., MOSIER J. N.: *Guidelines for designing user interface software*. Mitre Corporation, Bedford, MA, 1986. 6

[SS02] SEO J., SHNEIDERMAN B.: Interactively exploring hierarchical clustering results. *Computer 35*, 7 (2002), 80. 3

[SS05] SHAWKY H. A., SMITH D. M.: Optimal number of stock holdings in mutual fund portfolios based on market performance. *Financial Review 40*, 4 (2005), 481. 2, 3

[SS07] SANCETTA A., SATCHELL S. E.: Changing correlation and equity portfolio diversification failure for linear factor models during market declines. *Applied Mathematical Finance 14*, 3 (2007), 227. 4

[STZM11] SONG D.-M., TUMMINELLO M., ZHOU W.-X., MANTEGNA R. N.: Evolution of worldwide stock markets, correlation structure, and correlation-based graphs. *Phys.Rev.E 84*, 2 (2011), 026108. 4, 6

[TADMM05] TUMMINELLO M., ASTE T., DI MATTEO T., MANTEGNA R. N.: A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences of the United States of America 102*, 30 (2005), 10421–10426. 3

[TDMAM07] TUMMINELLO M., DI MATTEO T., ASTE T., MANTEGNA R. N.: Correlation based networks of equity returns sampled at different time horizons. *The European Physical Journal B 55*, 2 (2007), 209. 3

[TKBH17] TURKAY C., KAYA E., BALCISOY S., HAUSER H.: Designing progressive and interactive analytics processes for high-dimensional data analysis. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 131–140. 7, 10

[TLL10] TSE C. K., LIU J., LAU F. C. M.: A network perspective of the stock market. *Journal of Empirical Finance 17*, 4 (2010), 659. 3, 6

[TLM10] TUMMINELLO M., LILLO F., MANTEGNA R. N.: Correlation, hierarchies, and networks in financial markets. *Journal of Economic Behavior and Organization 75*, 1 (2010), 40. 3

[TMB02] TVERSKY B., MORRISON J. B., BETRANCOURT M.: Animation: can it facilitate? *International Journal of Human-Computer Studies 57*, 4 (2002), 247–262. 10

[TSL*17] TURKAY C., SLINGSBY A., LAHTINEN K., BUTT S., DYKES J.: Supporting theoretically-grounded model building in the social sciences through interactive visualisation. *Neurocomputing 268* (2017), 153–163. 3

[Wik18] WIKIPEDIA: Greek government-debt crisis — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Greek%20government-debt%20crisis&oldid=835895744, 2018. [Online; accessed 13-April-2018]. 8

[YPH*] YANG J., PATRO A., HUANG S., MEHTA N., WARD M. O., RUNDENSTEINER E. A.: Value and relation display for interactive exploration of high dimensional datasets. In *Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on*, IEEE, pp. 73–80. 3

[ZMM12] ZHANG Z., MCDONNELL K. T., MUELLER K.: A network-based interface for the exploration of high-dimensional data spaces. In *2012 IEEE Pacific Visualization Symposium* (2012), p. 17. 3, 6

[ZMZM15] ZHANG Z., MCDONNELL K. T., ZADOK E., MUELLER K.: Visual correlation analysis of numerical and categorical data on the correlation map. *IEEE Transactions on Visualization and Computer Graphics 21*, 2 (2015), 289. 3, 6