Volume xx (200y), Number z, pp. 1–12

iFUNDit: Visual Profiling of Fund Investment Styles

R. Zhang[1], B. K. Ku[2], Y. Wang[3], X. Yue[4], S. Liu[5], K. Li[6], and H. Qu[1]

[1]Hong Kong University of Science and Technology, Hong Kong
[2]LookinBody Co. Ltd, Korea
[3]Singapore Management University, Singapore
[4]Baiont Technology Co. Ltd., China
[5]Smeal College of Business, Penn State University, USA
[6]RiceQuant Co. Ltd., China

Abstract

Mutual funds are becoming increasingly popular with the emergence of internet finance. Clear profiling of a fund's investment style is crucial for fund managers to evaluate their investment strategies, and for investors to understand their investment. However, it is challenging to profile a fund's investment style as it requires a comprehensive analysis of complex multi-dimensional temporal data. In addition, different fund managers and investors have different focuses when analyzing a fund's investment style. To address the issue, we propose iFUNDit, an interactive visual analytic system for fund investment style analysis. The system decomposes a fund's critical features into performance attributes and investment style factors, and visualizes them in a set of coupled views: a fund and manager view, to delineate the distribution of funds' and managers' critical attributes on the market; a cluster view, to show the similarity of investment styles between different funds; and a detail view, to analyze the evolution of fund investment style. The system provides a holistic overview of fund data and facilitates a streamlined analysis of investment style at both the fund and the manager level. The effectiveness and usability of the system are demonstrated through domain expert interviews and case studies by using a real mutual fund dataset.

CCS Concepts

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

1. Introduction

A mutual fund is an investment tool that allows many investors to pool money together to purchase securities. Compared to direct investment in individual securities, a mutual fund has outstanding advantages in terms of professional investment management and risk diversification, making it a popular investment choice. For example, 44.6% of households in the United States invested in mutual funds in 2019 [Ins22]. In China, the mutual funds market size increased by 27% in 2020, reaching over 18 trillion Chinese Yuan [oChi20]. The global mutual fund market size was US$53.9 trillion in 2020, and is projected to reach US$101.2 trillion in 2027 [Aar20].

Investment style is the dominant principle for the classification and analysis of mutual funds. Clear profiling of investment style is crucial as it reveals a fund's underlying investment strategy that determines the fund's performance [FF10]. The investment style provides invaluable insights for fund managers to evaluate their investment strategies, and for investors to understand and assess their portfolios. However, it is challenging to analyze a fund's investment style. To capture the underlying investment strategy, analysts often need to evaluate a fund from a variety of aspects, such as the types of stocks, the economic sectors, and the trading frequency. This process generates a large amount of high-dimensional temporal data, which is often difficult to navigate even for experienced fund managers. Furthermore, different investors and fund managers focus on different attributes of a fund when conducting investment style analysis. As a result, there is no standard for investment style analysis in the mutual fund industry. Institutions often develop customized models and in-house systems to define and analyze the investment styles of funds. According to the interviews with domain experts we collaborated with in this study, fund managers often rely on dedicated departments of their institutes in practice to investigate the investment styles of competitor funds.

Due to the complexity of fund investment style analysis and the lack of standard tools, the evaluation of fund investment style is even more challenging for investors. As a consequence, investors often choose a fund based on a single attribute (such as return or fund size), its prospectus, or even its name. Unfortunately, none of these can reflect a fund's actual investment strategy nor support a proper investment decision. Sometimes investors bet on a fund solely based on its fund managers, who determine the fund's investment style. However, fund managers may adjust their investment styles over time. In addition, under some circumstances, for a fund that is under the names of multiple fund managers, a listed manager in the fund's prospectus may not be the one who determines the fund's actual investment styles, which makes it hard for investors to confirm the exact investment style of a mutual fund and probably leads to a deviation from the investor's expectation.

Owing to its significance, various methods of investment style analysis have been developed. Some studies focus on the qualitative classification of individual funds; while others conduct quantitative analysis based on specific data sets, such as return/risk, or fund stock holdings. Visualization techniques are also proposed to facilitate the analysis of individual funds or the networks of fund managers. A detailed review of these methods is presented in Section 2.

However, to conduct a comprehensive analysis of funds' investment styles, we need to evaluate both the performance metrics and the stock holdings. In addition, the bi-partite relations between funds and the managers, which are the key factors of fund investment strategies and performance, should be investigated. Furthermore, it is crucial to provide benchmarks such as the overall return/risk and investment style of the market for evaluation as benchmarks are important principles in the industry. Overall, the analysis process should be streamlined for investors to assess thoroughly.

To address the above challenges, we propose *iFUNDit*, a streamlined analytics system that provides a holistic overview of fund data for profiling and comparing the investment styles of mutual funds. The acronym stands for "Interactive FUND Investigation Tool". The system aims to assist fund managers and potentially experienced investors in conducting in-depth analysis on different features of mutual funds including the style factors that reflect funds' underlying trading strategies, and the performance metrics that measure the results of such strategies. It also supports an investigation into the bi-partite relations between funds and managers which provide insights into the evolution of fund investment styles. In particular, we incorporate the multi-factor Barra Risk model [GK00], the GICS economic sector categorization [MSC20], and critical performance metrics into *iFUNDit*. Further, we propose a set of coupled visualizations: a distribution view to delineate the distribution of critical attributes of funds and managers, a cluster view to present the investment-style crowdedness on the market, and a detailed view to visualize the evolution of funds' investment styles. We evaluated the effectiveness of *iFUNDit* through case studies with domain experts from multiple institution by using the China mutual fund dataset.

The major contributions of our study are summarized as follows:

- A visual presentation of fund investment style to visualize the detailed composition and evolution of investment styles, and enable efficient comparison between different investment styles.
- An interactive visual analytics system, to provide a holistic view of fund data and streamline the analysis of fund investment styles. The system visualizes benchmarks to enable statistical analysis of funds/managers, which is critical for financial data analysis.
- A set of comprehensive case studies and interviews with domain experts from various financial institutions.

2. Related Work

We present a literature review on fund investment style and financial data visualization in this section.

2.1. Fund Investment Style

Mutual fund investment styles are important signals for investors, as it directly attributes to the performances of different funds. Owing to its significance, monumental efforts have been put into developing different methods for investment style analysis. Depending on the nature of the description of the investment styles, the analysis is either qualitative or quantitative.

In a qualitative analysis, investigators analyze fund managers through their reports, speeches, or by conducting interviews with them in order to infer their investment styles. In the industry, people have adopted quasi-quantitative evaluation methods to categorize a fund's investment style. One widely accepted method has been developed by Morningstar, Inc. [Mor20]. It constructs a set of criteria for the analysis of an investment style. Based on the criteria, its analysts evaluate a fund and give a certain score, which eventually categorizes the fund into a three-by-three Style-Box matrix with a qualitative description. Compared to quantitative factor models, this method is more accessible to the public. It is convenient to qualitatively label a fund with a preset investment style, but this falls short concerning a detailed analysis of a fund's investment style.

Quantitative research interprets styles through various factor models. Sharp, et al. introduced the classic CAPM one-factor model[Sha64; Sha77] to evaluate the returns of stocks, based on which a return-based style analysis RBSA model was proposed [Sha88]. Fama and Fench et al. developed another classic three-factor Fama-Fench model to explain stock return[FF92; FF10]. The model was then utilized to infer a funds' investment style [Dav01]. Carhart et al. added a momentum factor to the Fama-Fench model; thus, proposing a four-factor model[Car97]. Recently, Bar Rosenberg developed a multi-factor model, referred to as the Barra Risk Factor Analysis model, to interpret stock returns from more dimensions [GK00; BBMS13]. Depending on specific markets, the Barra Risk model constructs a comprehensive set of factors to evaluate stock returns, and can also be used to analyze the investment style of a stock portfolio. The Barra Risk model has been widely used for this, and it has been updated for different markets. For example, in the China stock market, the CNE-5 model which contains 10 style factors was proposed by MSCI in 2012. Compared with other factor models, the Barra model interprets an investment style from more dimensions, which offers more explanatory options for fund managers and investors to portray an investment style. In this study, we adopted the CNE-5 model for investment style analysis.

Based on these models, people adopted return-based and holding-based approaches for fund investment style analysis [KS07; DW12]. The former approach built models based on funds' performance, such as return, volatility and etc. The latter focused on funds' stock holdings which involve the study of the economic sectors or the company fundamentals of the stocks. Researchers also proposed various measures to evaluate the managerial skills and investment styles of fund managers, such as Reliance on Public Information[KS07], Active Share [KNV14], time-varying skills [CP09], etc.

The above-mentioned methods are designed to analyze a single fund. In order to compare different funds or evaluate a fund in a different time period, visualization techniques are indispensable.

2.2. Financial Data Visualization

In this study, the fund investment style data is multivariate time-series data, which is composed of performance data, holding data, and sector data. General multivariate time-series techniques have been studied in the past decades for various applications [Pen08; SSGM18; SRJ*17; KPB14; WLSL16; YKS*19; HKA09].

For applications in the finance context, some studies focused on improving classic visual forms such as scatter plot, line chart, and matrix to contain more information. StockViz [Saw09] utilized scatter plots in spiral arrangement to visualize the historical stock prices of individual companies. Matthias et al.[SWK*11] used line charts with segmented background and color encoding to show stock returns. Ziegler et al.[ZJGK10] proposed a pixel-based performance matrix to visualize volatility and return of funds in long-term investment. Yue et al. [YBL*19] proposed a system to visualize the risk factors in portfolio management with radar charts and line charts. However, it only uses the return to measure the portfolio's performance and does not support the analysis of managers.

In order to increase the information contained in a two-dimensional space, visualization techniques in the form of more sophisticated shapes have been developed. A number of studies were based on heat map techniques [AZZ09; ZNK08; ZNK07]. Alsakran et al. proposed a density-based distribution map and tile-based parallel coordinate system [AZZ09] to visualize multivariate financial data[AZZ10]. Csallner et al. and Jungmeister et al. [JT92] adopted tree-map graphs to visualize the stock holding diversity of mutual funds [CHLS03]. Xiong et al. showed the performance of funds in geographic maps [XPH02]. Lei et al. analyzed the volatility of the stock market in a ring-shaped design [LZ10]. David et al. proposed a glyph-based framework for visualizing multivariate data that can be adopted for finance applications [CLP*15].

Comprehensive surveys about visualization techniques were conducted. Aigner et al. summarized time-series related visualization for general application [AMM*07]. Ko et al conducted a survey on visual analysis for financial data [KCA*16]. FinanceVis.net [DML14] summarized finance-related visualizations in an interactive system.

In this study, in addition to the investment style of funds, we also aim to visualize the bi-partite relations between funds and managers. Techniques regarding bipartite visualization were reported in various applications [ZXQ15; XCQS16; CXDR18]. However, these techniques are not suitable to handle time-series data and cannot be simply applied in our study.

3. Design Requirements

This section introduces the background and the task analysis.

3.1. Background

A consensus about investment is that there is no single "best" investment strategy. High return is often associated with high risk. The balance between the two is made based on investors' personal judgment and preferences. In mutual fund investment, different people use different metrics to evaluate funds/managers from different perspectives. According to our survey and the domain experts, there is no standard tool to characterize and evaluate mutual funds' investment styles. Fund managers often rely on dedicated analysts in their institutions to investigate and summarize the strategies of competitor funds, and then conduct further analysis which is time-consuming. Therefore, it is desirable to build a system to streamline the process of fund investment style analysis.

3.2. Task Analysis

To gain a comprehensive understanding of the fund industry, we collaborated with researchers and practitioners from multiple research and industrial institutions. In addition to the co-authors, we worked closely with 5 domain experts from 4 different financial institutions.

The expert E_1 is a senior fund manager with more than ten years of experience in a top national fund institution. He manages over \$10 billion assets. In his daily routine, he works with analysts to study the investment styles of funds on the market, which involves a lot of offline investigation. However, many activities are restricted during the pandemic lockdown. He is eager to have an analytics tool to boost his team's analysis efficiency. E_2, E_3, and E_4 are from a financial service provider that serves overall 300 financial institutions and 120,000 individual users. E_2 is a fund researcher. E_3 (a co-author) and E_4 are financial product managers. These experts have extensive experience with financial institutions and investors, and have a comprehensive understanding of industry requirements. Furthermore, we also consulted two stock traders, E_5 and E_6, from another two trading firms to have a better understanding of the stock market. These domain experts are not co-authors except for E_3, who provided invaluable data and insight for this study.

We followed a user-centered design framework in this study. To ensure that the system fulfills domain users' requirements, we conducted a series of structured interviews with the domain experts to understand the problems and identify their concerns. Through the seventeen-month collaboration, we conducted online meetings for discussion and prototype demonstration during the pandemic lockdown. After a number of discussions and system development iterations, we have extracted the design requirements and decompose them to a list of tasks on three levels: single-fund level, multiple-fund level and system level.

Single-fund level:

T.1 *Characterize and visualize the investment style of a fund.* It involves many attributes to profile a fund's investment style, such as the stock types, the economic sectors and the trading frequency. The complexity of these multi-dimensional data is difficult to interpret even for experienced fund managers. It is crucial to present these data in a lucid visual form so investors and managers can have a clear holistic overview of fund's investment style.

T.2 *Visualize the temporal evolution of the investment style of a fund.* The investment style of a fund is intrinsically dynamic. A fund manager may adjust the investment style periodically. The visualization should capture the dynamic changes of the investment style in different time periods.

T.3 *Evaluate the correlation between the investment style and performance of a fund.* It is crucial to evaluate a fund's performance, such as return and risk. These performance metrics are the results of an investment style, and often have a great influence on investors' final investment decision. The visualization should provide a clear mapping between a fund's investment style and its performance metrics, so that investors and fund managers can easily explore the correlation between them.

Multiple-fund level

T.4 *Discover similar/different funds, in terms of investment styles or performance metrics.* It is critical for investors to be able to select funds with different investment styles to diversify the investment risks. It is also beneficial for fund managers to group funds with similar investment styles together in order to study the styles systematically. The system should be able to cluster funds according to both the investment styles and performance metrics.

T.5 *Compare different investment styles efficiently.* Conventionally, it is difficult to compare different investment styles in detail, no matter in a qualitative or quantitative approach. In a qualitative approach, it is hard to compare two styles in detail due to the lack of quantitative attributes. In a quantitative approach, many attributes are involved to characterize an investment style. It is not efficient to compare these attributes one by one. To address the issue, the system should not only visualize an investment style in a detailed manner, but also enable an effective comparison between different styles.

T.6 *Compare both the performance metrics and investment style of different funds in the context of benchmarks.* Benchmarks, such as the overall distribution of the funds on market and the market index, are critical for evaluating financial products. It gives investors a sense of how good or bad a fund is. For example, comparing to the description of "A fund has a return of 30%.", it is more informative for decision-making to present that "A fund has a return that exceeds 75% of all funds in the market." .

T.7 *Visualize the bi-partite relations between funds and managers.* Clearly profiling of the relations helps investors and analysts to identify which manager dominates the investment style of a fund so as to evaluate the fund more comprehensively.

System level

T.8 *Enable users to explore different attributes for evaluating a fund.* Different users focus on different attributes when evaluating a fund. An investor may value more on the return, while a fund manager may pay closer attention to a certain investment style factor such as the capitalization size of the stocks in order to analyze the style from a certain perspective. The system should enable users to explore different attributes through intuitive interactions.

4. System Pipeline

iFUNDit is a web-based application that is comprised of three modules: the database module, the data-processing module and the visualization module. The database module employs MongoDB to store both the raw data from RQData API and the processed data. Raw data refers to unprocessed data that consists of full records of Chinese funds on asset allocation, asset values, holdings, financial indicators and manager records, and other related information such as daily stock prices and daily stock factor exposures. This study focuses on the analysis of securities investment funds, of which the performance is significantly dependent on the investment style. The dataset consists of 2398 securities investment funds. Processed data consists of temporal data at fund- and manager-level which is calculated from the raw data. It includes fund-level information on quarterly factor exposure, quarterly sector positions, daily financial indicators and daily asset values.In addition to fund- and manager-level information, processed data also consists of mapping information from fund to manager and vice versa.

The data-processing module, implemented with Python Pandas and Scikit-learn, handles data manipulation such as aggregation of temporal data, filtering, and clustering algorithm. The visualization module is used for communication of data to the users with carefully designed visualizations. The system consists of six interactive views that function in harmony to provide visual-assisted investigation on funds investment styles and performances.

5. Data Models

This section introduces the data and the investment style factors.

5.1. Investment Style Factors

In this study, we proposed to combine the 10 style factors from the Barra China Equity Model (CNE5) [MSC13] and 11 sector-factors from the Global Industry Classification Standard (GICS) [MSC20] to construct a 21-factor metrics to characterize a fund's investment style.

The Barra's Risk Model proposes a set of generalizable risk factors existing in the market. This state-of-the-art model has been widely adopted by the industry to measure the risk factors associated with a stock relative to the market. The CNE5 model is a Barra's Risk Model designed for the China market, on which the dataset in this study is based. Since the fundamentals of different Barra's Risk Models for different markets do not vary substantially, our style factors can be easily applied to the global market. GICS classification is also a global standard which means our investment style metrics can be easily generalized to different markets worldwide.

In the CNE5 model, there are in total 10 style factors to evaluate the investment style from different perspectives which help investors to align the risk model with their investment processes. The 10 factors are *Beta*, *Book to Price*, *Earning Yield*, *Growth*, *Leverage*, *Liquidity*, *Momentum*, *Non-linear Size*, *Residual Volatility*, and *Size*.

These factors are calculated to capture the short-term and long-term dynamics of the market. The factor model [GK13] uses an assumption that there exists a set of K common factors that drive stock returns. The equation for stock return can be written as: $r_i^t = \sum_{k=1}^{K} X_{ik} f_k^t + \varepsilon_i^t$, where $i = 1,2,\ldots,M$, $t = 1,2,\ldots,T$; r_i^t is the return of stock i at time t; f_k^t is the return of factor k at time t; X_{ik} is the factor exposure of stock i on factor k for the time period $t = 1,2,..T$; and ε_i^t is the specific return of stock i, which cannot be explained by the factors at time t. Based on the equation, multivariate linear regression [RC12] is used to estimate the factor exposure of a stock $\{X_{ik}\}_{i=1,2,\ldots,M;k=1,2,\ldots,K}$.

Consider a fund consisting of N stocks, a weight of stock i is w_i , then the return of this fund at time t is the weighted average of individual stock returns: $R_F^t = \sum_{i=1}^{N} w_i r_i^t$. The fund's exposure to factor k is given by the weighted average of the stock exposure: $X_k^F = \sum_{i=1}^{N} w_i X_{ik}$.

The style factor and factor exposure together reveal the types of stocks that a mutual fund holds, which reflects its investment style.

Factor exposure of the 10 style factors is used to measure how much a fund is exposed to a certain style factor, which infers the investment style of the fund. Positive factor exposure means the fund exposes to a certain style factor more than the market index, while negative factor exposure means the fund expose to a certain style factor less than the market index. Zero factor exposure means the fund has the same exposure as the market index to a certain style factor. For example, a positive *Size* factor exposure means the fund allocates more asset in stocks with large market capitalization, or “large-cap" in short, while a negative *Size* one means the fund prefers small-cap stocks. A high *Book to Price* factor exposure indicates that the fund implements a “value" investment strategy, which means it buys stocks whose market price is “underrated", while low *Book to Price* shows that the fund uses a “growth" strategy and invests in stocks that are “overpriced" but have potential to grow even bigger.

In addition to the 10 style factors, we constructed 11 sector factors based on the Global Industry Classification Standard (GICS), to characterize a fund’s investment style in detail. The sector factors indicate to which economic sector a stock belongs. We used one-hot encoding to indicate whether a stock belongs to a specific sector.

In summary, the 10 style factors and the 11 sector factors together construct the 21-factor metrics to profile funds’ investment styles.

5.2. Unsupervised Clustering Algorithms

To identify funds with similar/distinct investment styles (**T4**), we need to cluster funds according to their investment styles. We utilize two unsupervised dimensionality reduction and clustering algorithms, namely, t-SNE[MH08] and MDS[Kru64] to facilitate this task. The t-SNE method preserves the local proximity structure and creates tight clusters for visualization. It is useful to identify funds that have similar investment styles. MDS tends to retain the distance between clusters. Investors can use this method to seek funds with distinct investment styles so as to diversify the investment risks.

Parameter settings of the clustering algorithms can pose a significant impact on the results. However, the target users of our system often do not have a technical background. Therefore, to ensure clear clustering results, we have conducted experiments on our dataset and established empirical rules to set parameters dynamically according to the number of data points [WVJ16].

6. Visual Design

iFUNDit consists of six coordinated views: the Manager View (Fig. 1B), the Fund View (Fig. 1C), the Cluster View (Fig. 1D), the List View (Fig. 1E), the Detail View (Fig. 1F) and the Temporal View (Fig. 1G). We carefully design the system framework by following Shneiderman’s mantra [Shn03]: *“Overview first, zoom and filter, then details on demand”*.

The system starts with filtering the attributes that users want to focus on in the Control Panel (Figs. 1A). It then provides an overview of the performance metrics and style features of funds through the Manager View, the Fund View, and Cluster View (Figs. 1B,C,D). Users can brush the points of interest in any of the three views. Upon the brushing, the corresponding points in the other two views are connected with lines across the three views, as shown in Fig.1. At the meantime, details of the selected funds is visualized in the List View, the Detail View, and the Temporal View (Figs. 1E,F,G). Users sort the list of funds according to a selected attribute by clicking the corresponding column in the List View. Simultaneously, the rows of glyphs in the Detail View which represent the evolvement of funds’ investment style, and the curves in the List View which illustrate the performance of funds, are sorted accordingly.

We make use of benchmarks, and relative scales such as quartiles to provide references for evaluation. In the context of financial investment, evaluation often involves comparisons to some generalizable benchmarks or ranking to indicate whether a fund is *better or worse* instead of absolute *good or bad*. Therefore, in our system, the performance metrics and investment styles of funds are presented in a comparative context, which is critical for fund analysis.

6.1. Manager View and Fund View

The Manager View (Fig. 1B) and the Fund View (Fig. 1C) provides an overview of the distribution of all funds in two-dimensional space and facilitate exploration of the relative positions of funds and managers in the entire market (**T4, T6**). The two views are used in combination with the Cluster View to visualize the mapping between the performance and the investment styles (**T3**). The three views serve as the entry point of the analysis workflow.

An augmented scatter plot is used in the Manager View and the Fund View. The two views enable a clear assessment of a fund’s performance with regard to its competitors. Users can select what attributes to be represented on the two axes in the two views by adjusting the parameters in the Control Panel. Each axis of the scatter plot is equipped with a colored quartile ribbon, a density plot, and projections of selected funds ((Fig. 1B, C)). This helps users to understand the positions of the selected funds/managers in the distribution (**T6**), which is critical in the context of financial metrics. The color of points in the scatter plot encodes the latest net asset size managed by funds and managers. The vertical axis of the Manager and Fund View are aligned side-by-side to allow the mapping of values between the two views.

Users can select specific managers or funds by direct queries using the search box embedded in the Manager View and Fund View, or by brushing the points of interest in the Manager View, the Fund View, and the Cluster View. Upon selection, the corresponding points of a fund/manager across the three views are connected with polylines, which help users easily track the points across the three views. The corresponding manager(s) and fund(s) will also be added to the List View, the Detail View, and the Temporal View for single-fund level analysis.

Justification: We deliberately incorporate quartile ribbons, distribution density plots, and links in the Manager View and Fund View to facilitate the evaluation of funds and managers’ performance.

1) *Use of colored quartile ribbons.* An alternative method for visualizing quartile is box plot. However, overlaying box plot on the axes of the scatter plot can confuse the users because both plots use

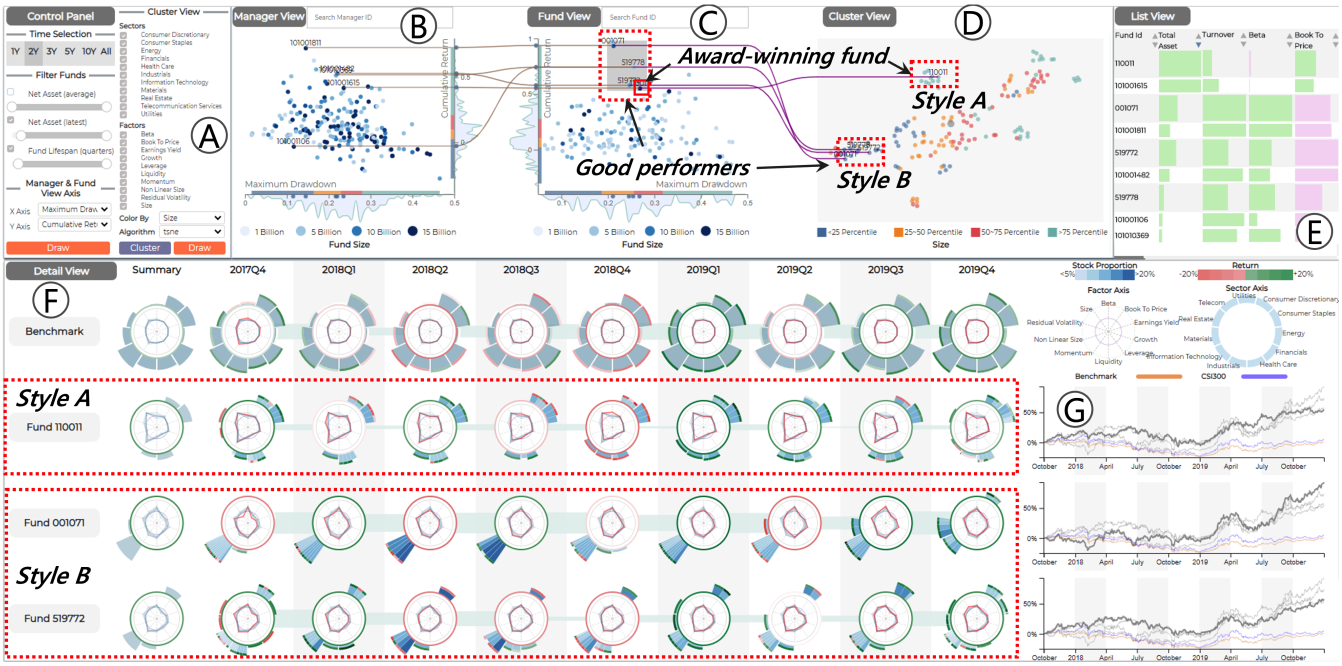

Figure 1: *iFUNDit provides a holistic view of fund performance metrics and investment style factors at the fund and the manager level. (A) Control Panel supports interactive exploration with different attributes that users focus on. (B) Manager View shows the distribution of fund managers' performance attributes, such as return and risk. (C) Fund View displays the distribution of funds' performance attributes. (D) Cluster View projects the crowdedness of fund investment styles. (E) List View lays out the selected funds and managers. (F) Detail View visualizes the evolution of funds' investment styles. (G) Temporal View displays the return of funds and benchmarks.*

shape lines as the visual channel. Hence, we use the color channel to encode quartile ranges which are more distinguishable. The color scheme is consistent with the quartile color scheme used in the Cluster View, minimizing the color diversity.

2) *Use of color to encode asset size of funds and managers.* In Manager View and Fund View, we use the color of a circle instead of its area to encode the asset size of a fund. This contradicts with the common visualization practice of using magnitudinal channel to encode quantitative values. This decision was made considering the nature of the fund market distribution, where the majority of the funds dwell near the "average", forming a distribution with high kurtosis. Using circles of varying sizes in these dense regions can result in severe visual clutter and make each circle indistinguishable.

3) *Use of links to connect associated funds and managers.* An alternative method to display connectivity between associated funds and managers would be highlighting each fund-manager pair using different colors. However, such design suffers from scalability issues, as the introduction of additional color encoding would make it difficult to distinguish different fund-manager pairs, especially when there are many funds/managers selected. In addition, it interferes with the color encoding in the Cluster View which represents the quantile of the selected attributes.

6.2. Cluster View

The Cluster View (Fig. 1D) visualizes the clustering of funds with respect to their investment styles. Multi-dimensional style factors are projected onto a 2D plane by using unsupervised clustering algorithms (Section 5.4).

In the investment style analysis, there is no standard classification which causes difficulties in identifying investment style clusters (**T4**). To address this issue, users can customize the features for clustering, and select one attribute to be color encoded in the view from the Control Panel (**T8**). Since the features used for color labels are numerical values, we transformed the numerical values into quantile categories which are widely used in the fund industry. Users can choose either t-SNE or MDS as the desired clustering algorithm.

6.3. List View

The List View (Fig. 1E) displays the details of other performance metrics of funds and managers that are not shown in the Manager View and the Fund View. The List View allows the comparison and sorting of the selected funds from various perspectives (**T6, T8**). The performance metrics of managers associated with the funds are also shown, enabling a convenient comparison of managers within a fund or across different funds (**T7**).

Each row represents a fund with aggregated values of various features. The magnitudes of values are encoded using horizontal bars, which allows an intuitive comparison across rows and columns. The color of the bars encodes the sign of the values (i.e. green for positive and red for negative). Users can sort the rows according to a desired attribute by clicking on the corresponding column. Rows are expandable and collapsible on click to display the associated managers of funds. The row arrangement of the lists view is synchronized with the order of funds in the Detail View. This allows different arrangements of funds for style comparison.

6.4. Detail View

The Detail View (Fig. 1F) visualizes the details of funds in terms of investment style, performance, and stock holdings (**T1**). The Detail View depicts the temporal evolution of a fund (**T2**). Through observation of the style and performance changes of a fund, users can estimate how the style changes have influenced the fund's performance (**T3**). The Detail View also provides insights into the rationale behind the style change of a fund, which can be caused by the change of managers or the market conditions. In addition to single-fund analysis, the Detail View enables in-depth comparisons of investment styles between funds (**T5**).

Each row shows a fund's investment style and performance over time. Each glyph encodes the style and performance of each quarter. The performance (quarter return) is represented in two granularity: fund-level and individual stock-level. The uppermost row in the Detail View displays the benchmark which is the average values of all stock-based funds operated in the time period, that allows users to observe the overall trend and performance of the market (**T6**).

Glyph Design The glyph consists of two main components: the *center radar chart* and the *circular sectors* as shown in Fig. 2. The *center radar chart* encodes 10 style-factors of a fund. The red line represents the current quarter and the light blue lines represent the other quarters. Since it is difficult to diagnose the differences from separated radar charts, we deliberately add radar charts of other quarters in each glyph so that they serve as direct references for comparison. The *colored ring* (Fig. 2a) between the radar chart and the circular sectors encodes the fund's overall quarterly return. *Circular sectors* (Fig. 2) encode various information of about 11 sectors. To avoid confusion and to emphasize that the two parts have distinct encoding, we add a gap in between to separate them, and intentionally misalign the axes of the two parts. Each circular sector (Fig. 2b) presents a sector factor and is fix-oriented in designated directions. Its height encodes the sector ratio. Each circular sector is partitioned, with each partition representing an individual stock holding. The size of each partition encodes the sector-level stock holding ratio. Each partition is divided into its inner (Fig. 2c) and outer part (Fig. 2d). The color of the inner part of the partition encodes the fund-level stock holding ratio. The color of the outer part of the partition encodes the quarterly return of the stocks. Hovering over the circular sectors displays the sector-specific stock details such as the stock ID, quarterly return, and holding ratio. Adjacent glyphs of each fund are connected by bridges (Fig. 2e), whose thickness encodes the turnover rates. The change of managers is marked by a circle (Fig. 2f) on the bridges. Hovering over the circle shows the details of the managers.

Justification: We make nontrivial efforts to incorporate domain-friendly visualizations such as pie charts and radar charts to minimize the difficulty of using the design. An alternative design to visualize the 11 sector factors is by using a donut chart instead of fix-oriented circular sectors. Different sector factors are encoded

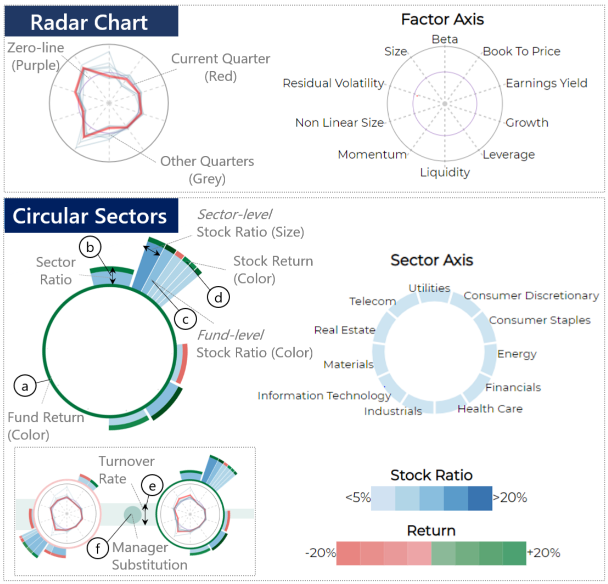

Figure 2: *The glyph consists of two main components. The radar chart (top) encodes 10 style-factor values of a fund. The circular sectors (bottom) encode information about 11 sector-factors. These two components visualize the 21 style features that characterize an investment style.*

using categorical colors. The size of the donut is proportional to the sector holding ratio. Given that the sector holding ratios always add up to 100 percent, a donut chart can be an effective visualization design for representing such proportional values. However, it is possible that two donut charts with totally different sector ratio can have indistinctive shape except colors. Moreover, frequent changes of the positions of each donut sector make it difficult to observe sector-factor-wise changes. Additionally, the donut chart does not encode zero values, which makes it difficult to spot empty sectors. Fix-oriented circular sectors overcome these issues by providing a clear visual distinction between sector changes and empty sectors. It also provides an additional visual channel for displaying the return of individual stocks encoded by the outer part of the circular sector.

6.5. Temporal View

The Temporal View (Fig. 1G) shows the evolution and the trend of selected funds' performance during the selected time period. The temporal pattern of the performance of each selected fund is plotted next to the fund's glyphs that are shown in the Detail View. The performance of each fund is highlighted with a dark grey curve in the plot in the corresponding row, while the other funds are shown in light grey in the same plot to achieve easy comparison. The benchmark and CSI300 Index are plotted in orange and purple, respectively. The benchmark shows the average return of funds on the market. The CSI300 is an important benchmark in the mutual fund industry to evaluate funds' performance [Com20].

By combining the Detail View and the Temporal View, *iFUNDit* establishes a link between the evolution of a fund's investment styles and its performance to facilitate in-depth analysis (**T3, T6**). On the one hand, users can use the Temporal View to assess the performance that is analyzed in the Detail View. On the other hand, after identifying an interested trend in the Temporal View, users can use the Detail View to analyze the associated investment style during this period. In addition, this view helps to identify temporal patterns of fund performance, and supports a direct performance comparison between funds and against the benchmarks(**T6**).

Justification: Line charts with multiple lines can bring severe visual clutters due to the dense crossing lines and color diversity, which makes it inefficient to explore the temporal evolution of different funds. To alleviate this issue, we allocate each fund with a dedicated plot and highlight its performance with dark grey color in the corresponding plot, rather than using a single plot to show all funds with no emphasis. By reading the Detail View and Temporal View horizontally, users can analyze the correlation between the investment style and the performance of a fund (**T3**). By comparing different plots in the Temporal View vertically, users can quickly identify different trends of different funds' performance (**T6**).

7. Case Studies

We present two case studies in this section to demonstrate how the domain experts E_1 - E_4 used the system to profile fund investment styles and analyze the fund managers. The notations, **T1 - T8**, are used to mark the tasks that are associated with the design requirements discussed in Section 3.2.

7.1. Fund-level investigation

What are the investment styles of good funds? The fund manager, E_1, would like to analyze the investment styles of top-performing funds on the market (**T1, T3**). This was a routine he performed regularly at work. Normally, the analysis was performed by a dedicated team in his fund institute. The team analyzed good-performing funds and competitor funds, and summarize reports to E_1.

E_1 investigated funds with the asset size of over \$3 Billion and with at least two-year history, and decided to study their two-year performances (**T8**). He selected the two axes for the Manager and Fund View as *cumulative return* and *maximum drawdown*, which are the two key performance attributes he prioritized (**T8**). E_1 then selected all 21 style features to cluster funds, and used the *Size* factor for the color label in the Cluster View (**T8**) as he wanted to focus on the capital size of stocks that the funds invest in, and used the default t-SNE clustering. E_1 set these parameters in the Control Panel (Fig.1A), and clicked the "Draw" button to plot the distributions of funds that satisfied his requirements.

E_1 observed several funds with relatively high return and medium risk in the distribution (**T6**), hinted by the quantile ribbons. He brushed the region in the Fund View (as shown in Fig.1B) to select the funds for detailed investigation (**T2**). Upon brushing, names of selected funds and managers were displayed with connecting links in the Fund View and the Cluster View (**T3**). From the name of the fund, E_1 recognizes Fund A, which received the *Fund-of-the-Year-2020 Award* from Morningstar[Mor20] in March 2020. E_1 then examined the Cluster View, and found that the investment styles fell in two distinct clusters, as shown in Fig.1D. From the relative positions and color labels, E_1 recognized that the two clusters have contrasting investment behavior in terms of *Size* factor (**T1, T5, T7**). The top cluster in Cyan color, belonged to the upper 75 percentile group of the *Size* factor, which inferred that the funds mainly invested in large-cap stocks (Style A). On the other hand, the bottom cluster in Blue color, belonged to the lower 25 percentile group which indicated that it invested in small-cap stock (Style B).

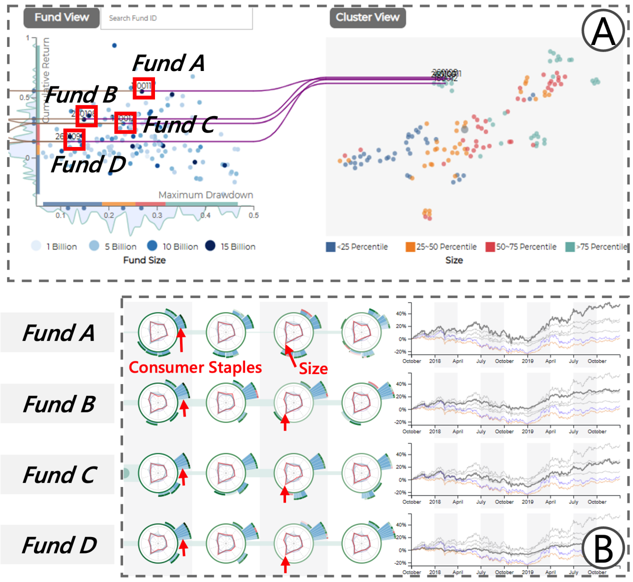

Figure 3: *Identifying funds that are similar to a given investment style. A) The user brushed the given investment style in the Cluster View, to show all funds with similar styles in the Fund View. B) The details of these funds were visualized in the Detailed View and Temporal View.*

After observing how Fund A separated from the others, E_1 decided to investigate what made Fund A special. He proceeded to the Detail View for more comprehensive analysis of investment styles (**T2**). He observed that the glyphs of Fund A showed clear differences from other funds in general (**T1, T5**). In particular, E_1 identified that Fund A focused on the *Consumer Staple* and *Consumer Discretionary* sector while other funds focused on the *Information Technology* and *Material* sector. E_1 mentioned that it was interesting to find that Fund A did not invest in the most popular Information Technology sector on the market, as shown by the Benchmark glyphs in the first row in the Detail View (**T6**). Instead, it pursued its own investment strategy. E_1 then read the turnover bridges and noticed that Fund A consistently had a low turnover rate which was indicated by the thin bridges, showing that Fund A pursued low-frequency trading unlike the other funds (**T1, T5**). E_1 then headed to the Temporal View, where he discovers that Fund A had always outperformed the CSI300 Index (purple) and the other selected funds also outperformed most of the time. E_1 also found that the market average (orange) always underperformed in terms of cumulative return in the last two years (**T2**). In order to review the other performance metrics of the funds, E_1 used the List View to sort the funds using various features. After inspecting at different angles, E_1 found that Fund A had significantly high assets under management and high historical return.

By summarizing the findings from *iFUNDit*, E_1 was able to conclude that Fund A pursued its own distinguishing investment strategy. It employed a low-frequency trading style and favored large-cap stocks in its unique sector selections. It was interesting to reveal that the Fund A's actual investment style was contradictory to what its name suggested, a "medium-to-small-cap" fund. E_1 was impressed that *iFUNDit* could help him to shortlist an awarding-winning fund from thousands of funds and also assisted him in profiling its investment style, with intuitive and effective interactions.

Looking for similar funds. E_1 was then curious about whether there were funds with similar investment styles to Fund A (**T4**) and how they performed. He brushed the Fund A's neighboring nodes in the Cluster View. Three other funds were highlighted in the Fund View as shown in Fig.3. E_1 evaluated their investment styles in the Detail View, and found that these funds had similar investment styles. The orientation of the circular sectors showed that these funds invested heavily in the *Consumer Staples* sector. The shape of the center radar charts revealed that these funds mainly invested in large-cap stocks, which is indicated by the sharp bulge in the direction of the *Size* factor axis.

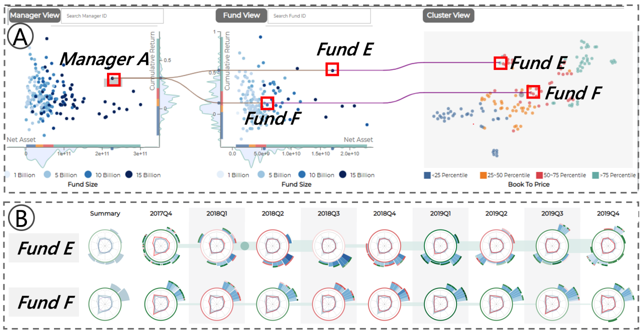

Figure 4: *Investigation of the style of a fund manager and his funds. A) A big-ticket fund manager had two funds with distinct investment styles and performance under his name. B) The Detail View showed the composition of the two funds' investment styles.*

Interestingly, E_1 said that he was well aware of the actual investment style of Fund B, because the manager of Fund B was once his colleague. E_1 confirmed that the actual style of the Fund B was indeed similar to that of Fund A. E_1 was impressed that *iFUNDit* could discover Fund B given Fund A. He mentioned that it normally requires a lot of domain knowledge and comprehensive investigation in order to identify similar funds with given criteria. He did not expect to accomplish the task with simple interactions in *iFUNDit*.

E_1 noticed that all the funds similar to Fund A had good performances. Their cumulative returns were all above 50% percentile of the market, and the downward risk (measured by maximum drawdown) were lower than 50% percentile of the market (**T4, T6**). The Temporal View showed that all funds consistently outperformed the CSI300 and the market average in the past two years. This finding gave E_1 a solid reference that could help him to adjust his own investment style.

7.2. Manager-level investigation

Fund manager is the most critical element that determines the investment style and performance of a fund. Investing in a fund is essentially betting on its fund managers. However, identifying the actual manager of a fund can be trickier than it sounds. Many funds have multiple managers, in which case the actual managers are difficult to identify. This is because fund institutes sometimes put their famous fund managers to the manager list of many funds, especially newly launched ones, to attract investors. However, these famous managers may not actually manage the fund. In practice, in order to identify whether a manager actually managers a fund, analysts gather information from various resources such as the manager's talks, news, and conduct interviews with the manager. The product managers E_2, E_3, and E_4, would like to look for clues about the actual manager of a fund (**T6**) to provide insights for their customers.

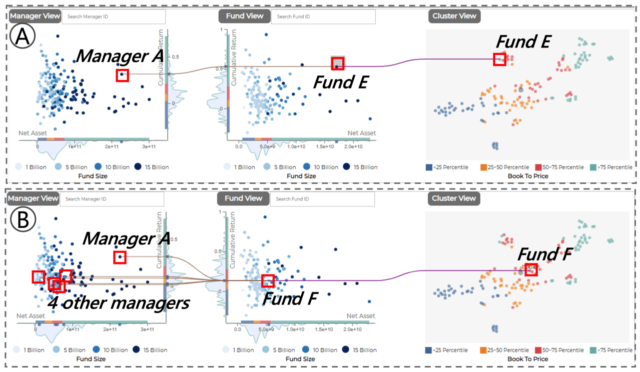

Figure 5: *Identify the actual manager. A) In Fund E, Manager A was the sole manager. B) In Fund F, Manager A was one of the five listed managers. Manager A's performance was distinct from the other four. It implied that Manager A was not the actual manager.*

E_2, E_3, and E_4 investigated on fund managers who manage funds with high net assets (**T7**). These managers are usually famous and assigned to the manager lists of many funds. E_4 set the "net asset" for the X-axis in the Manager and Fund View, and the "cumulative return" for the Y-axis (**T8**). With this setting, managers with large net assets were plotted towards the right side of the Manager View.

E_4 brushed a manager on the right side of the Manger View, whose cumulative return ranked above 75 percentile indicated by the top percentile ribbon on the Y-axis, to observe the funds under his names (**T7**), Fund E and Fund F, as shown in Fig.4. However, the two funds had different investment styles, indicated by their distinct positions in the Cluster View (**T4, T5**). This was confirmed in the Detail View. Fund E diversified its stocks in many sectors as indicated by the many non-empty circular sectors in the glyphs. On the contrary, Fund F almost invested solely in two sectors, which were not the focusing sectors of Fund E. In addition, Fund E had a high turnover rate (indicated by the thick bridge), and the investment style changed from time to time (indicated by the varying glyph shapes) (**T1, T2**), while Fund F had low turnover rate and maintained a consistent investment style during the period. These observations showed that Manager A was managing two funds with drastically different investment styles.

Typically it is unlikely for a fund manager to adopt drastically different investment styles at the same time. Therefore, E_4 suspected that Manager A was not the actual manager of both funds. To verify his hypothesis, he brushed the two funds to check their managers. The results are shown in Fig.5. Fund E had Manager A as its sole manager, which infers that its investment style was the actual investment style of Manager A. In contrast, Fund F had 5 managers. Considering the investment style of Fund F which is drastically different from Fund E, it was very likely that Manager A was not the actual manager of Fund F. It could be also observed from the Manager and Fund View that, Manager A and Fund E had greater return than the other managers and Fund F (**T6**). Therefore, E_2, E_3, and E_4 suspected that Manager A did not actually manage Fund F. The findings could help their customers make informed decisions.

8. Expert Interviews

It is crucial that the system is evaluated by domain experts to ensure it satisfies the design requirements as discussed in Section 3.2. We conducted interviews with ten domain experts to evaluate our system. In addition to the six experts introduced in Section 3.2, we further interview four. E_7 is the Chief Product Officer of a financial product institution who led the development of a state-of-the-art fund analysis tool for over ten leading financial institutions. E_8 is a trading strategist and quantitative analyst from a leading investment bank worldwide. E_9 is a fund manager of private equity. E_{10} is an experienced fund investor.

The interviews were conducted online due to the pandemic. Each interview lasted about 90 minutes. Firstly, we introduced the objectives of the interviews, explained the design, and demonstrated the system for 30 minutes. Then we let the experts explore the system for 30 minutes. Finally, we interviewed the experts for 30 minutes. The biggest challenge in the evaluation is that there is no standard tool or workflow for fund analysis, therefore it is difficult to set one unified benchmark to be compared with the system directly. To resolve that, we asked experts to compare the system with the tools they regularly use, and to give comments and suggestions. We summarize the feedback as follows.

Usability The experts appreciated that the workflow of *iFUNDit* is intuitive and efficient. E_1 stated that, compared to his regular tools, *iFUNDit* provided flexible starting points of analysis by separating the performance metrics and investment style factors in coordinated views. He could start his analysis from the performance metrics in the Fund View and Manager View, as well as from the investment styles in the Cluster View. These assessments are conducted with different tools in their daily practice and are more time-consuming. E_5 and E_7 suggested that the system could be too complex for regular investors who do not have sufficient domain knowledge. E_{10}, as an experienced investor, liked how the system allowed him to complete the tasks discussed in Section 3.2 without requiring him to refer to different resources.

Generalizability E_1 and E_8 mentioned that their institutes often define special performance metrics and factors to evaluate funds and managers, some of which are not included in the current system. It would be convenient for them to customize features. E_2 also confirmed that their customers often developed in-house models with customized features. To facilitate this requirement, *iFUNDit* allows users to define features for customized investment style analysis from the backend. The glyph design is capable of visualizing over twenty factors, which are sufficient for typical in-house models. Overall, the experts confirmed that the system can be applied to different markets and evaluation models by doing moderate adjustments to the investment style features and the performance metrics.

Visual Design Experts agreed that the design shows the details of investment styles clearly. They could inspect an investment style from different perspectives conveniently, and compare different styles efficiently. In their typical practice, the attributes that represent investment styles are presented in different forms separately. They appreciated *iFUNDit* connected different views through convenient interactions and clear visualization, which makes it more efficient to track attributes. They mentioned the glyph took them relatively more time to learn when introduced but they could understand the visual encoding easily. They appreciated that the Detail View showed the holistic evolvement of funds, and the glyph showed the investment sectors clearly, which was difficult to achieve with their regular tools. They appreciated that the system incorporated a novel glyph design together with conventional visualization such as scatter plots. Such a design reduces the learning curve which is crucial for fund managers. They appreciated how the system emphasizes benchmarks in all views. They particularly liked the percentile ribbon and density plot in the Manager View and Fund View. E_1 mentioned that, "*In industry, fund managers' KPI is often measured by their rankings in the market. The system efficiently shows the relative position of a fund or a manager in the market, which is crucial for our evaluation.* E_6 concerned about visual clutter in the Manager/Fund View. He suggested applying more filtering criteria to reduce the number of funds plotted at a time when starting the analysis.

Functionality Experts suggested additional functions that could be useful for users. E_7 mentioned that it would be convenient for fund managers to review funds if *iFUNDit* shows more textual information about funds such as media reports. E_8 suggested including the management fee of funds, which helps investors to select funds. Such information varies on different investment platforms. The information is not visualized in the current system, as it does not define the investment style of funds. Nevertheless, it is achievable to incorporate such information by means of adding separated view windows, or by showing it on hover-over tooltips.

9. Discussion

The case studies and expert interviews demonstrate the effectiveness of the system. Domain experts confirm that the system helps to profile fund investment style efficiently, and has a good potential to create impact in the industry. Although we used China mutual fund data in the study, the proposed analytics framework and the system are directly applicable to mutual funds in the global market, as well as to private funds if the data is accessible. Furthermore, the system can be applied to other evaluation models other than the Barra Risk Model as long as they have a similar number of factors.

The system has some limitations. One issue is the scalability of the Fund/Manager View. Due to the nature of the fund market, the performance attributes of many funds/managers dwell near the median of the overall distribution. This causes visual clutter in the two views, which can make it difficult to brush a certain fund/manager near the center of the distribution. In order to alleviate this issue, we used color to encode the circle size in the two views and implemented filters and direct queries to select funds/managers more easily. A zoom-in feature could also help users brush more precisely.

Another issue is the visibility of the glyph when the number of investment style factors becomes large. This is a common limitation in radial-layout visualization. To alleviate this issue, we select features on a proper aggregated level. For example, we use the 11 sector factor on the top aggregated level in the GICS industry classification on the glyph, and can present more detailed information using interaction techniques like how we display detailed stock information by hovering over a sector.

iFUNDit does not explicitly label a fund with a certain investment style. This is because a fund's investment style can be measured from different perspectives and interpreted differently. For example, a fund can be labeled as "large-cap" from the perspective of stock caps, or labeled as "value" from the perspective of the of stock value orientation. The objective of the study is to profile an investment style in detail, instead of labeling it qualitatively. Nevertheless, it would be convenient for users to have a certain label.

The system starts the analysis from the manager level and the fund level, then further goes down to the stock level. It does not support direct stock-level queries, such as finding funds that hold specific stocks or evaluating the number of stocks that two funds hold concurrently. The target users are mainly domain experts such as fund managers and fund advisors who can use the system to explain the investment styles of funds to investors. Experienced investors are also potential users but they need to have sufficient domain knowledge to use the system properly on their own.

10. Conclusion and Future Work

We investigate the visual analysis of mutual fund investment style. Through close collaboration with domain experts, we propose a visual analytics system, *iFUNDit*, to address two major challenges of fund investment style analysis: a detailed presentation of fund investment styles, and an effective approach to compare different styles. We categorize funds' critical attributes into two groups, namely, the performance metrics and the investment style factors, and visualize them cohesively to achieve a streamlined analysis of investment style at both the fund level and the manager level. The system emphasizes visualizing benchmarks so that the analysis can be conducted in a comparative context, which is critical for fund investment analysis. We conducted extensive evaluations including case studies and expert interviews by using a real mutual fund dataset to validate the usefulness and effectiveness of the system.

In the future, we plan to conduct research on characterizing the personalities of investors and fund managers to better guide fund investment and management in behavioral finance study [DdCGC19; Bel10]. In particular, we can use the investment style details obtained in the system to profile the personalities of fund managers and label them accordingly. We can integrate a Know-Your-Customer process, where the investment preference of investors could be profiled. With profiling of investors, funds, and managers, it is feasible to develop a recommendation system to match investors with funds according to investors' personality traits.

Acknowledgement

This research is partially supported by HK RGC GRF grant (No. 16210321).

References

[Aar20] Aarti G Pramod B, Vineet K. "Mutual fund assets market". (2020). https://www.alliedmarketresearch.com/mutual-fund-assets-market-A06932, last visit on 2023-5-5 1.

[AMM*07] Aigner, Wolfgang, Miksch, Silvia, Müller, Wolfgang, et al. "Visualizing time-oriented data—a systematic view". *Computers & Graphics* 31.3 (2007), 401–409 3.

[AZZ09] Alsakran, Jamal, Zhao, Ye, and Zhao, Xinlei. "Visual analysis of mutual fund performance". *2009 13th International Conference Information Visualisation*. IEEE. 2009, 252–259 3.

[AZZ10] Alsakran, Jamal, Zhao, Ye, and Zhao, Xinlei. "Tile-based parallel coordinates and its application in financial visualization". *Visualization and Data Analysis 2010*. Vol. 7530. International Society for Optics and Photonics. 2010, 753003 3.

[BBMS13] Bender, Jennifer, Briand, Remy, Melas, Dimitris, and Subramanian, Raman. "Foundations of factor investing". *https://www.msci.com* (2013). URL: https://www.msci.com/documents/1296102/1336482/Foundations_of_Factor_Investing.pdf 2.

[Bel10] Belcher, Lawrence J. "Prior perceptions, personality characteristics and portfolio preferences among fund managers: an experimental analysis". *Journal of Behavioral Finance* 11.4 (2010), 239–248 11.

[Car97] Carhart, Mark M. "On persistence in mutual fund performance". *The Journal of finance* 52.1 (1997), 57–82 2.

[CHLS03] Csallner, Christoph, Handte, Marcus, Lehmann, Othmar, and Stasko, John. "Fundexplorer: Supporting the diversification of mutual fund portfolios using context treemaps". *IEEE Symposium on Information Visualization 2003 (IEEE Cat. No. 03TH8714)*. IEEE. 2003, 203–208 3.

[CLP*15] Chung, David HS, Legg, Philip A, Parry, Matthew L, et al. "Glyph sorting: Interactive visualization for multi-dimensional data". *Information Visualization* 14.1 (2015), 76–90 3.

[Com20] Company, China Securities Index. "CSI300 Index". (2020) 7.

[CP09] Cremers, KJ Martijn and Petajisto, Antti. "How active is your fund manager? A new measure that predicts performance". *The review of financial studies* 22.9 (2009), 3329–3365 2.

[CXDR18] Chan, Gromit Yeuk-Yin, Xu, Panpan, Dai, Zeng, and Ren, Liu. "V i B r: Visualizing bipartite relations at scale with the minimum description length principle". *IEEE transactions on visualization and computer graphics* 25.1 (2018), 321–330 3.

[Dav01] Davis, James L. "Mutual fund performance and manager style". *Financial Analysts Journal* 57.1 (2001), 19–27 2.

[DdCGC19] De Bortoli, Daiane, da Costa Jr, Newton, Goulart, Marco, and Campara, Jéssica. "Personality traits and investor profile analysis: A behavioral finance study". *PloS one* 14.3 (2019) 11.

[DML14] Dumas, Maxime, McGuffin, Michael J, and Lemieux, Victoria L. "Financevis.net-a visual survey of financial data visualizations". *Poster Abstracts of IEEE Conference on Visualization*. Vol. 2. 2014 3.

[DW12] DING, BILL and WERMERS, RUSS. "Mutual fund performance and governance structure: The role of portfolio managers and boards of directors". *Available at SSRN 2207229* (2012) 2.

[FF10] FAMA, EUGENE F and FRENCH, KENNETH R. "Luck versus skill in the cross-section of mutual fund returns". *The journal of finance* 65.5 (2010), 1915–1947 1, 2.

[FF92] FAMA, EUGENE F and FRENCH, KENNETH R. "The cross-section of expected stock returns". *the Journal of Finance* 47.2 (1992), 427–465 2.

[GK00] GRINOLD, RICHARD C and KAHN, RONALD N. *Active Portfolio Management*. McGraw Hill New York, 2000, 73–185 2.

[GK13] GLANTZ, MORTON and KISSELL, ROBERT L. *Multi-asset risk modeling: techniques for a global economy in an electronic and algorithmic trading era*. Academic Press, 2013 4.

[HKA09] HEER, JEFFREY, KONG, NICHOLAS, and AGRAWALA, MANEESH. "Sizing the horizon: the effects of chart size and layering on the graphical perception of time series visualizations". *Proceedings of the SIGCHI conference on human factors in computing systems*. 2009, 1303–1312 3.

[Ins22] INSTITUTE, INVESTMENT COMPANY. "Investment Company Fact Book". (2022). https://www.icifactbook.org/pdf/2022_factbook.pdf, last visit on 2023-5-5, 117 1.

[JT92] JUNGMEISTER, WALTER-ALEXANDER and TURO, DAVID. *Adapting treemaps to stock portfolio visualization*. Tech. rep. 1992 3.

[KCA*16] KO, SUNGAHN, CHO, ISAAC, AFZAL, SHEHZAD, et al. "A survey on visual analysis approaches for financial data". *Computer Graphics Forum*. Vol. 35. 3. 2016, 599–617 3.

[KNV14] KACPERCZYK, MARCIN, NIEUWERBURGH, STIJN VAN, and VELDKAMP, LAURA. "Time-varying fund manager skill". *The Journal of Finance* 69.4 (2014), 1455–1484 2.

[KPB14] KRAUSE, JOSUA, PERER, ADAM, and BERTINI, ENRICO. "INFUSE: interactive feature selection for predictive modeling of high dimensional data". *IEEE transactions on visualization and computer graphics* 20.12 (2014), 1614–1623 3.

[Kru64] KRUSKAL, JOSEPH B. "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis". *Psychometrika* 29.1 (1964), 1–27 5.

[KS07] KACPERCZYK, MARCIN and SERU, AMIT. "Fund manager use of public information: New evidence on managerial skills". *The Journal of Finance* 62.2 (2007), 485–528 2.

[LZ10] LEI, SU TE and ZHANG, KANG. "A visual analytics system for financial time-series data". *Proceedings of the 3rd International Symposium on Visual Information Communication*. ACM. 2010, 20 3.

[MH08] MAATEN, LAURENS VAN DER and HINTON, GEOFFREY. "Visualizing data using t-SNE". *Journal of Machine Learning Research* 9.Nov (2008), 2579–2605 5.

[Mor20] MORNINGSTAR. "Morningstar Style Box". (2020). https://www.morningstar.com/content/dam/marketing/apac/au/pdfs/Legal/Stylebox_Factsheet.pdf, last visit on 2023-5-5 2, 8.

[MSC13] MSCI. "The Barra China Equity Model (CNE5)". (2013) 4.

[MSC20] MSCI. "Global Industry Classification Standard". (2020) 2, 4.

[oChi20] OF CHINA, ASSET MANAGEMENT ASSOCIATION. "Statistics of Mutual Fund". (2020). https://www.amac.org.cn/researchstatistics/datastatistics/mutualfundindustrydata/, last visit on 2023-5-5 1.

[Pen08] PENG, ROGER D. "A method for visualizing multivariate time series data". *Johns Hopkins University, Dept. of Biostatistics Working Papers, Working Paper 166* (2008). URL: https://biostats.bepress.com/jhubiostat/paper166/ 3.

[RC12] RENCHER, ALVIN C and CHRISTENSEN, WILLIAM F. "Chapter 10, Multivariate regression–Section 10.1, Introduction". *Methods of multivariate analysis, Wiley Series in Probability and Statistics* 709 (2012), 19 4.

[Saw09] SAWANT, AMIT. "StockViz: Analyzing the Trend of Stocks in Major Auto, Oil, Consumer, and Technology Companies." *MSV*. 2009, 278–284 3.

[Sha64] SHARPE, WILLIAM F. "Capital asset prices: A theory of market equilibrium under conditions of risk". *The journal of finance* 19.3 (1964), 425–442 2.

[Sha77] SHARPE, WILLIAM F. "The capital asset pricing model: a "multi-beta" interpretation". *Financial Dec Making Under Uncertainty*. Elsevier, 1977, 127–135 2.

[Sha88] SHARPE, WILLIAM F. "Determining a Fund's Effective Asset Mix, Investment Management Review". *December* 59 (1988), 69 2.

[Shn03] SHNEIDERMAN, BEN. "The eyes have it: A task by data type taxonomy for information visualizations". *The Craft of Information Visualization*. Elsevier, 2003, 364–371 5.

[SRJ*17] SANSEN, JORIS, RICHER, GAËLLE, JOURDE, TIMOTHÉE, et al. "Visual exploration of large multidimensional data using parallel coordinates on big data infrastructure". *Informatics*. Vol. 4. 3. Multidisciplinary Digital Publishing Institute. 2017, 21 3.

[SSGM18] STOPAR, LUKA, SKRABA, PRIMOZ, GROBELNIK, MARKO, and MLADENIC, DUNJA. "Streamstory: exploring multivariate time series on multiple scales". *IEEE transactions on visualization and computer graphics* 25.4 (2018), 1788–1802 3.

[SWK*11] SCHAEFER, MATTHIAS, WANNER, FRANZ, KAHL, ROMAN, et al. "A novel explorative visualization tool for financial time series data analysis". *VAW2 011: The Third International UKVAC Workshop on Visual Analytics*. 2011 3.

[WLSL16] WANG, JUNPENG, LIU, XIAOTONG, SHEN, HAN-WEI, and LIN, GUANG. "Multi-resolution climate ensemble parameter analysis with nested parallel coordinates plots". *IEEE transactions on visualization and computer graphics* 23.1 (2016), 81–90 3.

[WVJ16] WATTENBERG, MARTIN, VIÉGAS, FERNANDA, and JOHNSON, IAN. "How to Use t-SNE Effectively". *Distill* (2016). DOI: 10.23915/distill.00002. URL: http://distill.pub/2016/misread-tsne 5.

[XCQS16] XU, PANPAN, CAO, NAN, QU, HUAMIN, and STASKO, JOHN. "Interactive visual co-cluster analysis of bipartite graphs". *2016 IEEE Pacific Visualization Symposium (PacificVis)*. IEEE. 2016, 32–39 3.

[XPH02] XIONG, F, PRAKASH, EC, and HO, KW. "ER modeling and visualization of large mutual fund data". *Journal of Visualization* 5.2 (2002), 197–204 3.

[YBL*19] YUE, XUANWU, BAI, JIAXIN, LIU, QINHAN, et al. "sPortfolio: Stratified Visual Analysis of Stock Portfolios". Vol. 26. 1. IEEE, 2019, 601–610 3.

[YKS*19] YAU, CALVIN, KARIMZADEH, MORTEZA, SURAKITBANHARN, CHITTAYONG, et al. "Bridging the Data Analysis Communication Gap Utilizing a Three-Component Summarized Line Graph". *Computer Graphics Forum*. Vol. 38. 3. Wiley Online Library. 2019, 375–386 3.

[ZJGK10] ZIEGLER, HARTMUT, JENNY, MARCO, GRUSE, TINO, and KEIM, DANIEL A. "Visual market sector analysis for financial time series data". *2010 IEEE Symposium on Visual Analytics Science and Technology*. IEEE. 2010, 83–90 3.

[ZNK07] ZIEGLER, HARTMUT, NIETZSCHMANN, TILO, and KEIM, DANIEL A. "Relevance driven visualization of financial performance measures". *EuroVis 2007*. 2007, 19–26 3.

[ZNK08] ZIEGLER, HARTMUT, NIETZSCHMANN, TILO, and KEIM, DANIEL A. "Visual analytics on the financial market: Pixel-based analysis and comparison of long-term investments". *2008 12th International Conference Information Visualisation*. IEEE. 2008, 287–295 3.

[ZXQ15] ZHOU, HONG, XU, PANPAN, and QU, HUAMIN. "Visualization of bipartite relations between graphs and sets". *Journal of Visualization* 18.2 (2015), 159–172 3.