FMLens: Towards Better Scaffolding the Process of Fund Manager Selection in Fund Investments

Longfei Chen, Chen Cheng, He Wang, Xiyuan Wang, Yun Tian, Xuanwu Yue, Wong Kam-Kwai, Haipeng Zhang, Suting Hong, and Quan Li

Abstract—The fund investment industry heavily relies on the expertise of fund managers, who bear the responsibility of managing portfolios on behalf of clients. With their investment knowledge and professional skills, fund managers gain a competitive advantage over the average investor in the market. Consequently, investors prefer entrusting their investments to fund managers rather than directly investing in funds. For these investors, the primary concern is selecting a suitable fund manager. While previous studies have employed quantitative or qualitative methods to analyze various aspects of fund managers, such as performance metrics, personal characteristics, and performance persistence, they often face challenges when dealing with a large candidate space. Moreover, distinguishing whether a fund manager's performance stems from skill or luck poses a challenge, making it difficult to align with investors' preferences in the selection process. To address these challenges, this study characterizes the requirements of investors in selecting suitable fund managers and proposes an interactive visual analytics system called *FMLens*. This system streamlines the fund manager selection process, allowing investors to efficiently assess and deconstruct fund managers' investment styles and abilities across multiple dimensions. Additionally, the system empowers investors to scrutinize and compare fund managers' performances. The effectiveness of the approach is demonstrated through two case studies and a qualitative user study. Feedback from domain experts indicates that the system excels in analyzing fund managers from diverse perspectives, enhancing the efficiency of fund manager evaluation and selection.

Index Terms—Financial Data, Fund Manager Selection, Visual Analytics.

1 Introduction

With the expansion of financial markets, funds have emerged as a popular investment avenue. The performance of these funds is intricately tied to their overseers, known as fund managers, who bear the responsibility of handling investment portfolios on behalf of both individual and institutional clients. Possessing specialized investment knowledge and professional skills, fund managers enjoy a distinct advantage over general investors, often resulting in superior returns. As the fund market continues to grow, the significance of fund managers becomes increasingly evident, prompting more investors to prefer investing in fund managers rather than directly in funds [1]. This shift can be elucidated through three key factors: *(I) Consistent performance.* Selecting fund managers with a proven track record allows investors to anticipate a replication of past successes in managing other funds, leading to sustained performance improvement over time, even in cases of new funds or those with shorter track record [2], [3]. *(II) Consistency of investment philosophy.* A fund manager's investment philosophy, which typically includes their approach to risk management, research, and decision-making, significantly influences a fund's performance [4]. Investors who aligns with the fund manager's philosophy may prioritize the manager over the fund. *(III) Reducing the impact of short-term fluctuations.* Emphasizing the fund manager over the fund mitigates the impact of short-term fluctuations, enabling investors to rely on the fund manager's expertise in navigating volatile markets and making strategic adjustments over time.

Conventional approaches to assessing and selecting fund managers typically involve performance measurement [5], [6], persistence analysis [3], [7], and manager characteristics [4], [8]. While these approaches have exhibited preliminary effectiveness in evaluating fund managers in previous studies, they encounter various challenges that impede their capacity to support a comprehensive decision-making process for fund manager selection. First, the challenge arises from **a large pool of candidates characterized by diverse and heterogeneous attributes.** Specific fund managers possess a range of attributes, including personal background details, industry investment ratios, and performance metrics. Ideally, selecting fund managers from such a diverse pool should integrate various relevant data to provide investors with comprehensive information. However, the process of gathering, analyzing, and comparing information about each fund manager can be time-consuming and laborious. Moreover, biased propaganda from institutions

- *Longfei Chen, Chen Cheng, He Wang, Xiyuan Wang, Haipeng Zhang, and Quan Li (corresponding author) are with the School of Information Science and Technology, ShanghaiTech University, Shanghai Engineering Research Center of Intelligent Vision and Imaging, China, and the Key Laboratory of Intelligent Perception and Human-Machine Collaboration (ShanghaiTech University), Ministry of Education. E-mail: chenlf,chengchen,wanghe1,wangxy7,zhanghp,liquan@shanghaitech.edu.cn.*
- *Yun Tian is with CFETS Information Technology (Shanghai) Co., Ltd., China. E-mail: tianyun_zh@chinamoney.com.cn.*
- *Xuanwu Yue is with Baiont Technology and Sinovation Ventures AI Institute. E-mail: xuanwu.yue@connect.ust.hk.*
- *Wong Kam-Kwai is with the Department of Computer Science and Engineering, The Hong Kong University of Science and Technology. E-mail: kkwongar@connect.ust.hk.*
- *Suting Hong is with the School of Entrepreneurship and Management, ShanghaiTech University. E-mail: hongst@shanghaitech.edu.cn.*

Manuscript received June 19, 2021; revised April 19, 2022.

can sway investors' judgment, emphasizing the importance of grasping available information to avoid improper fund manager selection and potential jeopardy to future profits. Second, **the difficulty lies in distinguishing between luck from skill in a fund manager's performance.** Given the uncertainty of financial markets, a fund manager's performance over time often results from a combination of skill and luck [9]. Although extreme returns of many funds may be fortuitous, they are significant enough to impact their managers' overall performance. In essence, some fund managers may outperform over a period of time, but this does not guarantee the same level of performance in the future. Thus, it becomes essential to differentiate between skill and luck when evaluating fund managers. Third, **the opacity of fund managers' actions poses a challenge.** Despite extensive disclosure requirements, mutual fund investors cannot observe all actions of fund managers [10]. When a fund manager's investment strategy and decision-making process lack transparency, investors find it challenging to assess the fund's risk level, influencing their choice and judgment of the fund manager. Last, **the diversification of selection preferences adds complexity.** Investors have specific preferences in selecting fund managers, such as a preference for high-tech industries or consideration of annualized returns. These preferences are highly individualized, necessitating an interactive mechanism to aid in constructing a diverse selection process. Overall, selecting fund managers involves navigating multiple factors and balancing their relevance and importance to make informed decisions aligned with investors' preferences and objectives.

In the current market, tools for analyzing fund managers generally fall into three main categories: financial data terminals like *Wind* and *Bloomberg*, online trading platforms such as *Alipay* and *Tiantian Fund*, and rating agency websites like *Morningstar*. While these tools are designed to provide investors with extensive data on fund managers for comprehensive analysis, they primarily offer information at an individual level. This limitation poses a challenge in conducting cross-sectional comparisons between fund managers, thereby impeding the selection process. To address these challenges and better support investors in selecting fund managers, we propose a systematic approach. We initiate the process by outlining the essential requirements of investors during fund manager selection. Subsequently, we introduce a visual analytics system, named *FMLens*, designed to enhance the selection process for investors. Specifically, we encode information about fund managers as feature vectors and transform the candidate space into the feature space. We then deconstruct the fund manager's investment style from multiple dimensions, including investment diversification and performance metrics. Meanwhile, we also use financial data to simulate the opaque actions of fund managers. By providing various observation perspectives, investors can more effectively distinguish between skill and luck, understand fund managers' investment strategies and decision-making processes, and more accurately assess their true investment abilities. Finally, we develop an interactive selection workflow that takes into account investors' personal preferences, offering a comprehensive and personalized approach to the fund manager selection process. The major contributions of our study can be summarized as follows.

- We highlight limitations in current market tools for analyzing fund managers, emphasizing their tendency to provide information at an individual level and the resultant challenge in conducting cross-sectional comparisons between fund managers.
- We introduce the *FMLens* visual analytics system as a solution to the identified challenges, which integrates multidimensional information to support investors in selecting fund managers.
- We demonstrate the efficacy of our approach through two case studies, expert interviews, and a qualitative user study.

2 Related Work

2.1 Fund Manager Performance Analysis

In recent times, substantial attention has been directed towards scrutinizing the performance of fund managers by both scholars and practitioners. To assess fund manager performance, widely adopted quantitative approaches include the Sharpe ratio [11], Jensen's alpha [12], and the Sortino ratio [13]. Additionally, researchers have explored qualitative characteristics, such as the fund manager's experience, education, gender, and investment philosophy, to understand their impact on performance [4], [8], [14]. William F. Sharpe's style analysis [6] has been utilized to identify fund managers with consistent investment styles and strategies. Introducing the concept of "Active Share", Cremers et al. [15] found that mutual funds with higher "Active Share" tend to outperform benchmarks, suggesting that increased active management correlates with superior performance. A similar conclusion was reached by Kacperczyk et al. [10], who investigated fund managers' unobserved actions. Furthermore, Fama et al. [9] examined the impact of both luck and skill on fund manager performance and highlighted the challenges associated with differentiating between the two.

Numerous studies have delved into exploring the persistence of fund managers' performance. For instance, in a cross-sectional regression analysis covering a sample of 2086 funds from 1992 to 1999, Klaas P. Baks [2] calculated Jensen's alpha for each period, concluding that there is persistence in fund manager performance. Similarly, Berk et al. [3] introduced a novel measure of fund performance and found that cross-sectional differences persisted over the subsequent ten years, indicating the presence of long-term persistence in fund manager performance.

Informed by previous research, our system design integrates various factors influencing fund manager performance, such as investment preferences, performance metrics, and personal characteristics. Furthermore, we provide investors with the capability to track shifts in fund managers' performance metrics over time, facilitating a retrospective analysis of their historical fund management activities to assess the consistency of their performance.

2.2 Financial Data Visualization

The realm of financial data visualization has gained significant attention, as evidenced by various studies [16], [17],

[18], [19], [20]. Visualization methods, including line graphs, icons, and tree diagrams, play a crucial role in applications such as stock trading and quantitative investment analysis. For instance, Schaefer et al. [21] proposed a line graph visualization method that efficiently displays a large amount of non-overlapping data and line graphs by utilizing background colors based on a specific code. Ziegler et al. [22] employed pixel bar charts to provide real-time comparisons of stock, industry sector, and country/region returns and volatilities. These studies commonly leverage line charts to extract additional information. Sorenson et al. [23] creatively combined continuous line graphs with icon-based event representations to present financial data effectively. Treemaps [24] and dendrograms [25] are utilized for displaying stock holdings. In the realm of quantitative trading, sPortfolio [26] classifies and integrates investment strategies based on their characteristics, offering an interactive design for strategy selection and verification within large-scale investment strategies. Another work, iQUANT [27], employs an embedded glyph design to assist stock traders in selecting promising investment portfolios. Building on this, RankFIRST [28] enhances iQUANT by introducing interpretable factor calculation models and displaying factor rankings through a hierarchical slope graph design.

When it comes to fund visualization, FundExplorer [25] introduces a distorted tree diagram to visualize a fund's portfolio composition and remaining stocks, facilitating the identification of additional funds for portfolio diversification. Dwyer [29] proposed a 2.5-D graph to illustrate the results of fund clustering based on stock holdings. Notably, existing research in visual analytics has predominantly focused on representing stock data or quantitative trading, leaving fund manager data comparatively understudied. Recognizing this gap, we propose a novel approach grounded in interactive visualization techniques to empower investors in efficiently evaluating, comparing, and ultimately selecting fund managers.

2.3 Multivariate Time-Series Data Visualization

Fund manager data encompass time-series with various attributes, including, but not limited to, manager background details, industry investment ratios, and income-related information over time. Numerous studies have delved into the analysis of time-series data [30], [31], [32]. For instance, Lei et al. [33] introduced the concept of "visual signature" for financial time-series data, aiding professional analysts in promptly extracting valuable visual insights. Sorenson et al. [23] devised a scalable visual representation to describe numerous discrete timestamped events, catering to the needs of hundreds of thousands of financial market professionals.

Conversely, other notable visualization methods center on analyzing time-series data with multiple attribute changes, primarily focusing on designing efficient visualizations for detecting trends and patterns [34], [35], [36], [37], [38]. In our study, we integrate innovative visualization glyphs with a timeline design to portray multivariate attributes, illustrating the evolution of various characteristics of fund managers. Moreover, we provide target users with interactive features to compare fund managers across different time periods or events, facilitating a comprehensive examination of their performance.

3 Observational Study

3.1 Needfinding Interview

In order to gain practical insights into the needs of fund investors, we collaborated with eight experts possessing a minimum of five years of experience in the field of fund investment. These experts were sourced from two channels: (1) the financial technology and investment program at a partner university, and (2) a collaborating financial institution. Prior to the interviews, we secured informed consent from the domain experts, allowing us to disclose their detailed information, including age, gender, occupation, and years of experience (Yrs Exp), as outlined in Table 1. The interviews, lasting approximately one hour each, were carried out through Zoom, an online conferencing platform. The entire interview process was audio-recorded with the explicit consent of the participants.

TABLE 1: The demographic characteristics of the eight domain experts (identified as E1 through E8) who participated in the needfinding interview were collected and analyzed.

ID	Gender	Age	Occupation	Yrs Exp
E1	Male	45	Senior Investment Manager	17
E2	Female	38	Investment Major, Professor	11
E3	Male	33	FinTech Major, Professor	10
E4	Male	36	Financial Analyst	9
E5	Male	34	Fund Risk Analyst	7
E6	Female	28	Financial Analyst	6
E7	Female	25	FinTech Major, PhD	5
E8	Male	27	Investment Major, PhD	5

In the interview sessions, our initial inquiry sought validation of the experts' recognition of the significance of selecting fund managers within their real investment procedures. Upon affirmation, we proceeded to inquire about the prevalent methods they employ to choose fund managers from the market. Following this, we prompted the experts to elucidate their individual analytical approaches to the fund manager selection process. In the third phase, we delved into the analytical tools regularly utilized by the experts in their daily activities, aiming to uncover any associated drawbacks or inconveniences. Collaborating with the experts, we engaged in a brainstorming session to pinpoint reasonable enhancements for the identified issues. Finally, we meticulously organized the gathered data and initiated an iterative coding process [39] for a comprehensive analysis. Based on the needfinding interviews, several qualitative findings pertaining to fund investors were identified.

F.1 Fund managers play a crucial role in investments. Experts consistently emphasize the crucial role fund managers play in investment decisions. E2 articulated this by stating, "*the fund manager drives the performance, so I pay more attention to their track record and investment philosophy than the fund itself*". Moreover, several experts provided instances where choosing the right fund manager significantly contributed to superior investment outcomes. For instance, E4 recounted an experience of selecting a relatively obscure fund based on trust in the fund manager's expertise, leading to outstanding fund performance over time.

F.2 Challenges in navigating a diverse pool of fund managers: Experts highlighted the complexities associated with efficiently accessing and consolidating information about a broad spectrum of fund manager candidates. They pointed out the limitations of existing tools in offering a comprehensive snapshot of the candidate space, resulting in a substantial time investment during the initial research phase. E1 expressed, *"existing tools don't provide a comprehensive overview,"* and E6 proposed that *"a more streamlined approach would save time and enable us to focus on what really matters—the analysis."*

F.3 Diverse approaches in assessing fund managers. Investors employ a varied approach to evaluate fund managers, with different individuals prioritizing specific factors during the screening process. E5, for example, places a premium on the fund manager's risk control ability, while E7 embraces a *"high risk, high return"* investment strategy. Despite these differences, most experts endorse a comprehensive assessment that considers multiple factors such as performance, investment philosophy, and risk management. This ensures a thorough understanding of fund managers' strengths and weaknesses, as underscored by E4.

F.4 Existing tools lack the flexibility to tailor selection strategies to individual preferences. Experts consistentlly emphasized the crucial alignment between an investor's philosophy and a fund manager's approach. The need for tools allowing customization of selection strategies to prioritize personally significant factors for selecting fund managers was a recurring theme. E6 expressed, *I feel more secure in my investments and achieve better returns when my investment objectives align with the fund manager's approach."* Despite this importance, experts observed a lack of flexibility in existing tools for tailoring selection strategies. E8 suggested, *"an ideal system would enable me to rank fund managers based on my investment preferences, allowing me to focus on the factors that matter most to me."*

F.5 Evaluation of fund managers should go beyond metrics. Experts underscored the limitations of metrics in evaluating fund managers. The need to go beyond aggregate metrics was emphasized, particularly for fund managers overseeing multiple funds simultaneously. E3 highlighted the significance of scrutinizing claims of high annualized returns, stating, *"It is crucial to investigate if such returns are achieved across all managed funds or if they are limited to only one fund, thereby concealing poor performance in other funds."*

F.6 Understanding "Unobserved Actions" for insightful evaluation: The term "unobserved actions" includes confidential decisions, strategies, or behaviors employed by managers to adjust fund positions[1] [10]. Despite their potential impact on fund performance, limited data or lack of disclosure often hinder a comprehensive assessment of these actions. Current tools inadequately shed light on these "unobserved actions", resulting in an incomplete understanding for investors. E1 emphasized that these actions reflect a fund manager's responsiveness to market events, underscoring the necessity for improved tools to assess such actions effectively.

F.7 Necessity of Comparative Analysis for Fund Managers: E4 and E6 highlighted the importance of conducting comparisons during the evaluation of fund managers. They emphasized that existing tools lack adequate support for this type of analysis. E6 specifically mentioned that current tools only integrate fund manager data without providing sufficient support for effective comparisons. This limitation impedes their work and creates challenges in making well-informed decisions.

1. Fund position refers to the proportion of equity assets held by the fund as a percentage of the fund's total assets.

3.2 Requirement Characterization

After several rounds of discussion and ideation, the subsequent requirements (**R.1–R.6**) were deduced from the needfinding survey.

R.1 Provide an overview of the fund manager candidate pool. The first requirement addresses the necessity for an effective way to present investors with an overview of the existing fund manager candidate pool (**F.2**). The overview should include details about the investment style and performance of the fund managers. The objective is to reduce the information collection effort and empower investors to promptly and efficiently assess the fund manager landscape. A clear and concise visualization of this information can assist investors in identifying managers who meet their investment criteria.

R.2 Assess a fund manager's investment expertise across various dimensions. The second requirement involves the need to assess a fund manager's investment skills across multiple dimensions, including risk management, asset allocation, and market timing. To support this evaluation, detailed metrics should be provided, empowering investors to make well-informed investment decisions aligned with their individual preferences and risk tolerance, as highlighted by the finding of **F.3**.

R.3 Customize ranking strategies according to investor preferences. Every investor possesses unique goals and preferences, and these factors should be taken into account when ranking fund managers (**F.4**). Investors should be enabled to assign weights to their specific preferences, such as risk tolerance and returns. This customization enables generating personalized rankings that prioritize fund managers aligning with the investor's preferences. This approach ensures a more precise match between the investor and the selected fund manager.

R.4 Provide a comprehensive historical review of management records. The fourth requirement underscores the need for a thorough historical review of management records of fund managers, extending beyond the consideration of performance metrics alone (**F.5**). In cases where an individual oversees multiple funds, relying solely on performance metrics may not provide a complete picture of their abilities. To address this, we should provide investors with access to detailed historical management records, empowering them to conduct a more holistic analysis of the fund manager's abilities.

R.5 Simulate fund managers' position adjustment actions. As highlighted by E1, unobserved actions, i.e., position adjustment actions, reflect the fund manager's investment ability and style, playing a pivotal role in the investors' decision-making process (**F.6**). Consequently, we should incorporate a simulation feature for fund position changes. This functionality allows investors to delve into the fund manager's investment style, market judgment, and timing, providing insights that contribute to well-informed investment decisions.

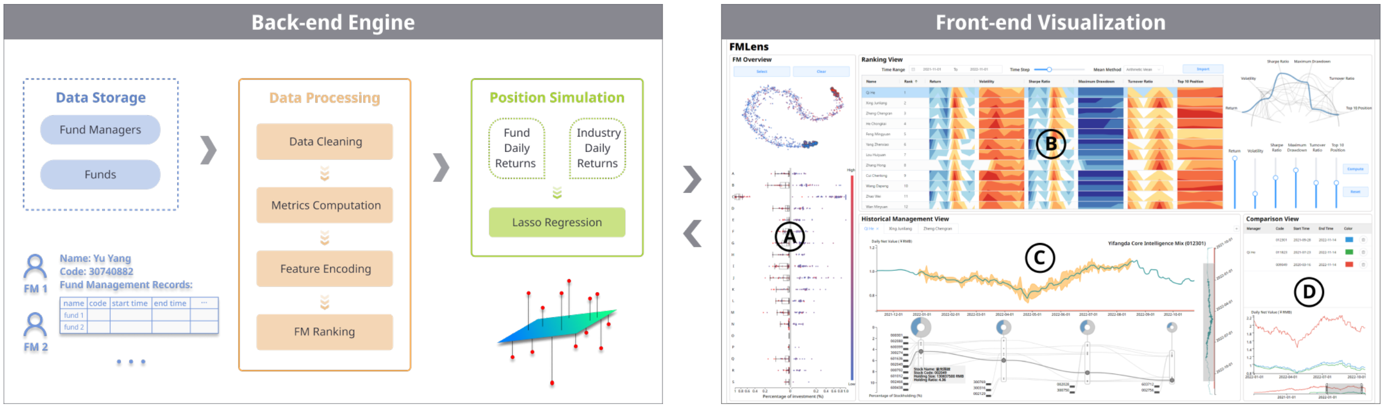


Fig. 1: The architecture of *FMLens* is composed of a back-end engine and a front-end visualization. Within the back-end engine, a data processing module converts the attributes of fund managers into performance metrics, and a simulation module employs regression techniques to predict adjustments made by fund managers in their positions. On the front-end visualization side, four distinct views have been carefully crafted for analysis.

R.6 Facilitate performance comparisons among fund managers. In line with the insights from experts (**F.7**), a crucial requirement is to enable comprehensive comparisons between fund managers, extending beyond a mere multi-level performance summary. Experts emphasize the significance of comparing the performance of products managed by fund managers over the same time frame. As articulated by E2, "*it is not uncommon for fund managers to yield similar high returns, yet for different reasons.*" Therefore, the ability to identify and understand these distinctions aids investors in making more informed decisions.

4 Approach Overview

We propose *FMLens*, an interactive visual analytics system that scaffolds the process of fund manager selection for investors. Figure 1 shows the architecture of *FMLens*, which consists of two primary components: a back-end engine and a front-end visualization. In the back-end engine, relevant attributes indicative of the fund managers' investment capabilities are initially quantified, and high-dimensional performance metrics are derived from raw data utilizing the data processing module. Subsequently, the simulation module employs regression techniques focused on the fund's net value to model the positional adjustments made by fund managers. In terms of front-end visualization, we have meticulously designed four distinct views aimed at providing comprehensive insights and facilitating interactive exploration of data. During the design process, we focused on user experience, undergoing multiple iterations of improvement to ensure that investors of varying financial and analytical expertise can utilize the system effectively.

5 Back-end Engine

5.1 Data Description and Processing

This study draws upon data obtained from a publicly accessible investment website[2]. As shown in Table 2, we filtered out equity fund managers from this dataset, resulting in a total of 2047 fund managers. Each entry in this dataset corresponds to a distinct fund manager and includes essential information such as the fund manager's *code*, *name*, *employment duration*, and *fund management records*. The latter encompasses the manager's historical fund management experience, providing details such as the fund's *code*, *name*, *management start time*, and *end time*.

TABLE 2: Two datasets used in our study.

Dataset	Capacity	Characteristics Dimensions
Fund Manager	2047	*Code*, *Name*, *Fund Management Records*
Fund	5439	*Code*, *Name*, *Size*, *Industry Allocation*, *Daily Net Value*, *Turnover Ratio*, *Top 10 Positions*

To accurately calculate fund manager performance metrics, we obtained additional information about the fund through the fund code using a publicly available data API[3]. Addressing the potential issue of duplicate queries arising from a fund being managed by multiple fund managers at different times, we first aggregated all funds appearing in the management records. Subsequently, we queried and stored data from all 5439 funds. Each entry in this resultant dataset corresponds to a unique fund and encompasses various informative variables, including *Daily net value*, *Size*, *Industry Allocation*, *Turnover Ratio*, and *Top 10 Positions*. It's noteworthy that the *daily net value* data is time-series in nature, updated at the conclusion of each trading day, while the other data points are discrete and disclosed quarterly in the fund's financial reports.

5.2 Performance Metrics Computation

Drawing from both relevant literature [3], [40] and expert recommendations, we have identified six key performance metrics for evaluating the investment style and abilities of fund managers:

- ***Return***: Also known as financial return, this metric denotes the net profit or loss generated by an investment within a specified time frame.

2. http://fund.eastmoney.com

3. https://www.ricequant.com/doc/rqdata/python/fund-mod.html

- ***Volatility***: A statistical measure reflecting the extent of variation in returns of a given security or market index.
- ***Sharpe Ratio***: This metric compares the return of an investment with its risk level, using a mathematical expression that considers the potential impact of excess returns over time, signaling greater volatility and risk than investment skill.
- ***Maximum Drawdown***: The maximum observed loss from the peak to the trough of a portfolio before a new peak is reached, serving as an indicator of downside risk over a given period.
- ***Turnover Ratio***: Representing the proportion of holdings within a mutual fund or other portfolios that are replaced within a given time period.
- ***Top 10 Positions***: Pertaining to the top ten stocks held within the fund's position, expressed as a percentage of net value.

It is important to note that the evaluation of a fund manager's performance evaluation is exclusively tied to the funds under their management. This implies that the previously mentioned metrics are computed on a per-fund basis [3], [12], [40]. For instance, consider a scenario where a fund manager oversees two funds, A and B. In the past three months, fund A has yielded a return of 20%, while fund B has produced a return of 40%. In this case, the fund manager's overall return is the average (30%) of the returns from both funds, irrespective of their size.

As suggested by experts, the presentation of the fund manager's performance metrics adopts a time-series perspective. Based on the characteristics of the data source behind the performance metrics, we categorize them into two types: "arbitrary interval" metrics and "fixed interval" metrics. The "arbitrary interval" metrics (including *Return*, *Volatility*, and *Sharpe Ratio*) are calculated based on the fund's net value of for each trading day, which is publicly available information. Therefore, we can specify any time interval to calculate the values of the three metrics as follows.

$$\textit{Return} = \sum_{k=1}^{n} w_k \times \frac{P_k^{(j)} - P_k^{(i)}}{P_k^{(i)}}, \tag{1}$$

where n represents the number of funds managed by the fund manager in the interval, $P_k^{(i)}$ denotes the net value of the k^{th} fund at the beginning of the interval, $P_k^{(j)}$ denotes the net value of the k^{th} fund at the end of the interval, and w_k represents the weighting, which is dependent on the averaging method;

$$\textit{Volatility} = \sum_{k=1}^{n} w_k \times \sigma_k, \tag{2}$$

where σ_k represents the standard deviation of the daily net value of the k^{th} fund within the interval; and

$$\textit{Sharpe Ratio} = \sum_{k=1}^{n} w_k \times \frac{R_p^{(k)} - R_f^{(k)}}{\sigma_p^{(k)}}, \tag{3}$$

where $R_p^{(k)}$ denotes the return of the k^{th} fund within the interval, $R_f^{(k)}$ denotes the risk-free rate of return of the k^{th} fund, and $\sigma_p^{(k)}$ denotes the standard deviation of the excess return of the k^{th} fund.

Unlike the "arbitrary interval" metrics, the "fixed interval" metrics (including *Maximum Drawdown*, *Turnover Ratio*, and *Top 10 Positions*) cannot be calculated using publicly available data; they must be directly obtained from the fund's quarterly reports. Since these reports are disclosed every three months, it limits these metrics to only reflecting the situation within fixed time intervals.

5.3 Fund Manager Ranking

Upon computing the performance metrics, we employ a TOPSIS [41]-based multi-attribute ranking approach. This approach considers the six aforementioned performance metrics as distinct attributes, aligning with expert recommendations. The ranking process unfolds through the following steps:

1) **Determining the Decision Matrix:** Each candidate manager is represented as a decision matrix D, where D_{ij} denotes the j^{th} attribute value of the i^{th} candidate.
2) **Standardizing the Decision Matrix:** Normalize the decision matrix D to yield the standardized decision matrix R.
3) **Determining Attribute Weights and Constructing the Weighted Matrix:** Utilize vector W to denote attribute weights set by the user, where w_j signifies the weight of the j^{th} attribute. Subsequently, construct the weighted standardized matrix V, with $V_{ij} = w_j \times R_{ij}$.
4) **Determining the Best and Worst Solutions:** Calculate the best solution A^+, where $A_j^+ = \max_i V_{ij}$, and the worst solution A^-, where $A_j^- = \min_i V_{ij}$.
5) **Computing the L^2 Distance between Each Candidate and the Best/Worst Solutions:** Calculate the distance $S_i^+ = \sqrt{\sum_{j=1}^{6}(V_{ij} - A_j^+)^2}$ from each candidate to the best solution and the distance $S_i^- = \sqrt{\sum_{j=1}^{6}(V_{ij} - A_j^-)^2}$ from each candidate to the worst solution using Euclidean distance.
6) **Ranking the Candidates:** Rank the candidates based on the similarity indices $c_i = \frac{S_i^-}{S_i^+ + S_i^-}$, where $0 \leq c_i \leq 1$. The condition $c_i = 1$ holds if and only if the candidate is the best solution, and $c_i = 0$ holds if and only if the candidate is the worst one. In essence, higher indices correspond to higher rankings, signifying superior choices.

5.4 Position Adjustment Simulation

As discussed in subsection 3.2 (**R.5**), simulating fund positions provides investors with a valuable exploration into the fund manager's investment style, market judgment, and timing. This exploration yields insights crucial for making well-informed investment decisions.

Generally, simulations of adjustments in fund positions fall into two primary categories: regression-based [42], [43], [44] approaches and Deep Learning (DL)-based approaches [45], [46]. Despite the robust capabilities of DL-based methods, they require extensive training data and often face computational efficiency challenges, limiting their real-time interaction. Consequently, we have chosen an index-based regression approach to model daily changes in fund positions—a method supported by various studies [42], [47], [48]. This methodology involves regression, with the daily returns of the fund as the dependent variable and the daily returns of primary industry indices[4] as independent variables. The problem can be formulated using the following regression equation:

$$R_{f,t} = \sum_{i=1}^{n} \gamma_{i,t} I_{i,t} + \varepsilon_t \quad (4)$$

In this equation, $R_{f,t}$ represents the daily returns of fund f on the t-th day, $I_{i,t}$ represents the daily returns of the i^{th} industry index on the t^{th} day, $\gamma_{i,t}$ is the regression coefficient to be fitted, and ε_t is the residual term. We interpret $\gamma_{i,t}$ as the percentage of the fund's investment in stocks belonging to the i^{th} industry on the t^{th} day. Consequently, $\sum_i \gamma_{i,t}$ represents the percentage of all assets held by the fund in equity assets, reflecting the fund position for the t^{th} day. The daily change in fund position is then measured by $\sum_i (\gamma_{i,t+1} - \gamma_{i,t})^2$.

To improve the fidelity of our simulations, we conducted a series of experiments to identify the optimal regression strategy and time window length that closely mimic real-world scenarios.

Regression Strategy: We initially employed three commonly utilized regression methods – Principal Component Regression (PCR) [49], Ridge Regression [50], and Lasso Regression [51]. These strategies were applied to estimate positions for 90 equity funds from the end of the third quarter of 2020 to the end of the third quarter of 2021, encompassing five quarter-end cross-sections. To simulate a fund's position on trading day T, we utilized daily return data for the fund and 28 primary industry indices over a span of 30 trading days, ranging from T-14 to T+15. Subsequently, we computed the Mean Absolute Error (MAE) [52] between the simulated and actual positions of the 90 funds on each quarterly reporting date. The results, presented in Table 3, indicate that Lasso Regression exhibits superior accuracy, Principal Component Regression demonstrates slightly weaker performance, and Ridge Regression consistently shows systematic overestimation.

TABLE 3: The obtained experimental results involve the application of three distinct regression strategies to simulate positions. For each regression method, we calculate the Mean Absolute Error (MAE) between the simulated and actual positions for 90 funds on each quarterly reporting date spanning from October 1, 2020, to December 31, 2021.

Reporting Date	PCR	Ridge	Lasso
2020/12/31	4.48%	3.54%	2.00%
2021/03/31	4.63%	4.15%	0.93%
2021/06/30	5.69%	9.94%	1.22%
2021/09/30	7.65%	12.98%	2.46%
2021/12/31	0.47%	13.39%	3.12%

Time window length. We further evaluate the sensitivity of the three regression methods to varying time window lengths, focusing on three cross-sections at the conclusion of the fourth quarter of 2020, the first quarter of 2021, and the second quarter of 2021, while varying window lengths from 15 to 59 days (Figure 2). Our observations reveal that, in most instances, the mean prediction error of each method stabilizes after the window length surpasses 40 days, signifying the convergence of solutions. No distinct pattern is discernible for window lengths less than 40 days. It is advisable to avoid excessively long window lengths, generally not exceeding one quarter (about 60 trading days), as the regression coefficients depict the average fund position in the past window period, and this value is employed to predict the fund position at the present time. An overly extended window length may result in lagging predictions. Based on our assessment, we opt for Lasso regression with a 30-day time window for our simulation module to align with the intensity and frequency of changes in fund positions.

4. A financial index produces a numeric score based on inputs such as a variety of asset prices. It can be used to track the performance of a group of assets in a standardized way. Indices typically measure the performance of a basket of securities intended to replicate a certain area of the market.

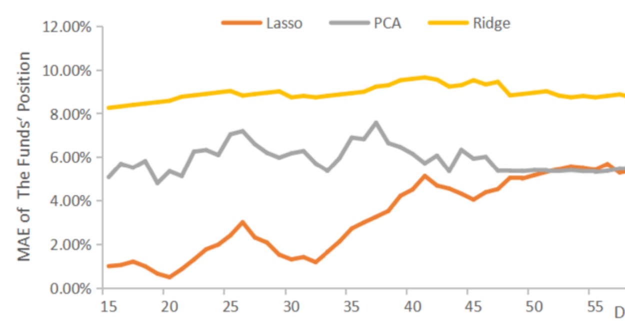


Fig. 2: The present chart displays the variation in the average mean of the length of the time window across three quarters, spanning a range of 15 to 59 days. Notably, the performance of Lasso regression stands out as the best, reaching its nadir at approximately 30 days.

6 FRONT-END VISUALIZATION

The fundamental design principle guiding *FMLens* aims to leverage and enhance familiar visual metaphors, allowing investors to focus on analysis. Illustrated in Figure 3, we have developed four primary visualizations to assist investors in the efficient evaluation and selection of potential fund managers: the *FM Overview* (Figure 3 (A)), summarizing the fund manager candidate space (**R.1**); the *Ranking View* (Figure 3 (B)), displaying the evolution of various performance metrics over time and facilitating fund manager ranking based on investor preferences (**R.2-R.3**); the *Historical Management View* (Figure 3 (C)), integrating past fund management records and simulating their positions to provide insights into the investment styles and abilities of fund managers (**R.4-R.5**); and the *Comparison View* (Figure 3 (D)), enabling investors to conduct a detailed comparison of the performance of different fund managers (**R.6**).

6.1 FM Overview

The *FM Overview* (Figure 3 (A)) serves as a summary of the candidate space, aiding in the identification of potential similarities and outliers among fund managers. Common dimensionality reduction techniques, including *t-SNE*, *PCA*, and *MDS*, are widely employed to create low-dimensional

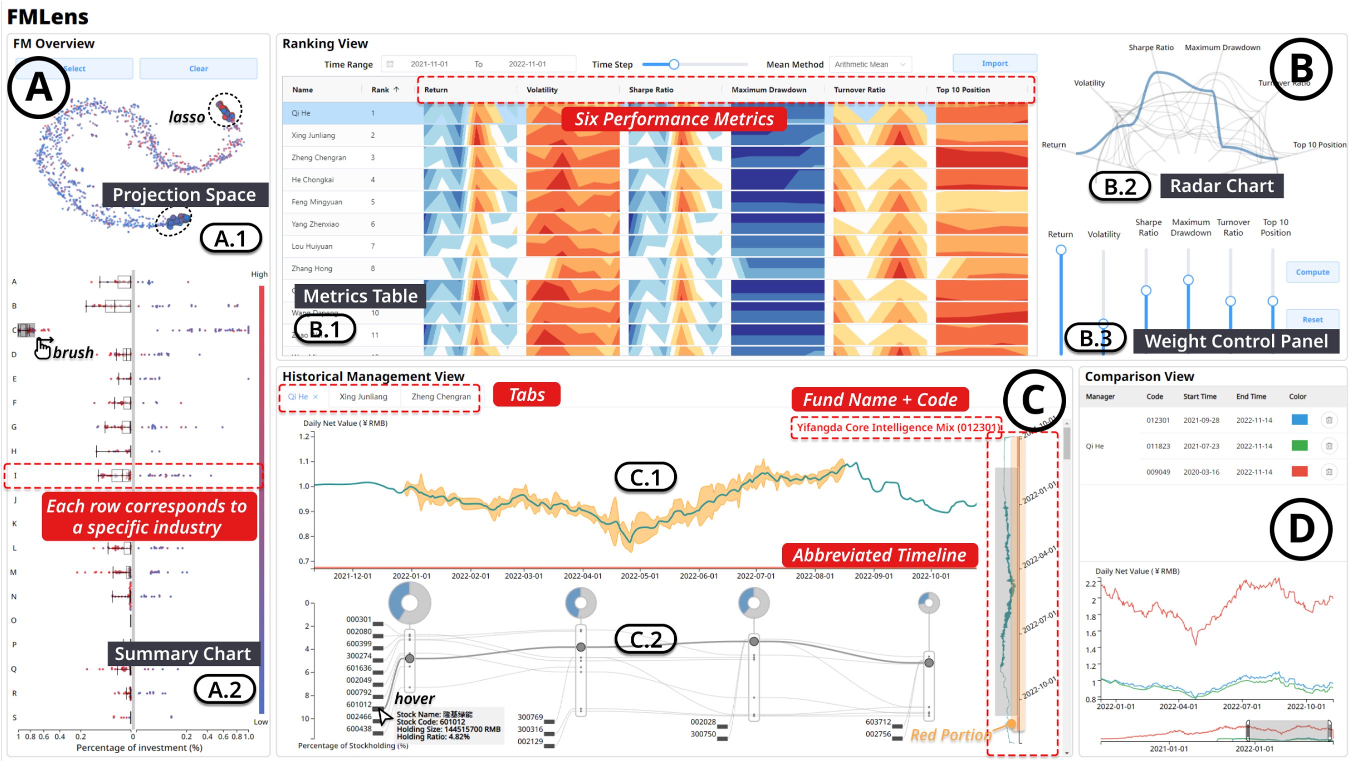


Fig. 3: The front-end visualizations of *FMLens* consists of four views: (A) The *FM Overview* serves as a summary of the fund manager candidate space. (B) The *Ranking View* facilitates the examination of fund managers' performance evolution and supports interactive ranking. (C) The *Historical Management View* provides a comprehensive review of fund managers' management records. (D) The *Comparison View* is crafted to facilitate the comparison of fund performance among one of more fund managers.

representations that preserve local similarities, revealing neighborhood structures [53], [54]. Adhering to conventional practices, we project all fund managers into a 2D space to explore potential clusters and outliers. Following discussions with domain experts, we utilize the amount of money invested by fund managers in each industry sector as feature dimensions to assess their investment styles at the industry level. Specifically, we quantify each fund manager's cumulative investment amount in the 19 primary industry sectors defined by the China Securities Regulatory Commission (CSRS) as the feature vector. We applied three dimensionality reduction methods—*PCA*, *MDS*, and *t-SNE*—to scrutinize the data. The results indicated that *PCA* and *MDS* methods yielded suboptimal performance, while the *t-SNE* method consistently generated projection results aligned along a curve. In consultation with experts, it was proposed that this tendency could be attributed to the nature of high-dimensional data, potentially reflecting correlations or inter-industry relationships influencing the fund manager's investment decisions and manifesting as a curve in the low-dimensional space. Subsequently, we utilized *t-SNE* to project these high-dimensional vectors into a 2D space. As illustrated in Figure 3 (A.1), each point in the projection space represents a fund manager, with color indicating cumulative returns.

Users can lasso any of the points and click the "Select" button to display the specific industry allocation of the lassoed fund managers in the summary chart below. Due to space constraints, our system supports selecting at most two groups, which will be displayed on the left and right sides of the summary chart. Users can also clear all selected points by clicking the "Clear" button. In Figure 3 (A.2), the summary chart consists of 19 rows corresponding to the 19 primary industry sectors mentioned earlier, with letters (**A**–**S**) on the left representing the industry codes defined by the CSRS. For a lassoed fund manager, each row features a point representing the person's investment percentage in that industry, with the color of the point corresponding to the person's return in that industry, maintaining consistency with the projection space. Users can brush a region in any row, prompting the system to select all fund managers in that region as a group and draw a box plot in each row, illustrating the distribution of the group's investment percentage in each industry.

6.2 Ranking View

The *Ranking View* (Figure 3 (B)) streamlines the examination of the evolution and overall distribution of fund managers' performance metrics over time, offering the capability to assign weights to these metrics for ranking purposes. This view is composed of three subviews: the metrics table (Figure 3 (B.1)), the semicircular radar chart (Figure 3 (B.2)), and the ranking control panel (Figure 3 (B.3)).

Metrics Table. Traditional financial software commonly uses line graphs to illustrate the dynamic changes in metrics over time. However, the goal of this study is to enable a more efficient comparison of dynamic metric changes among various fund managers within a confined space.

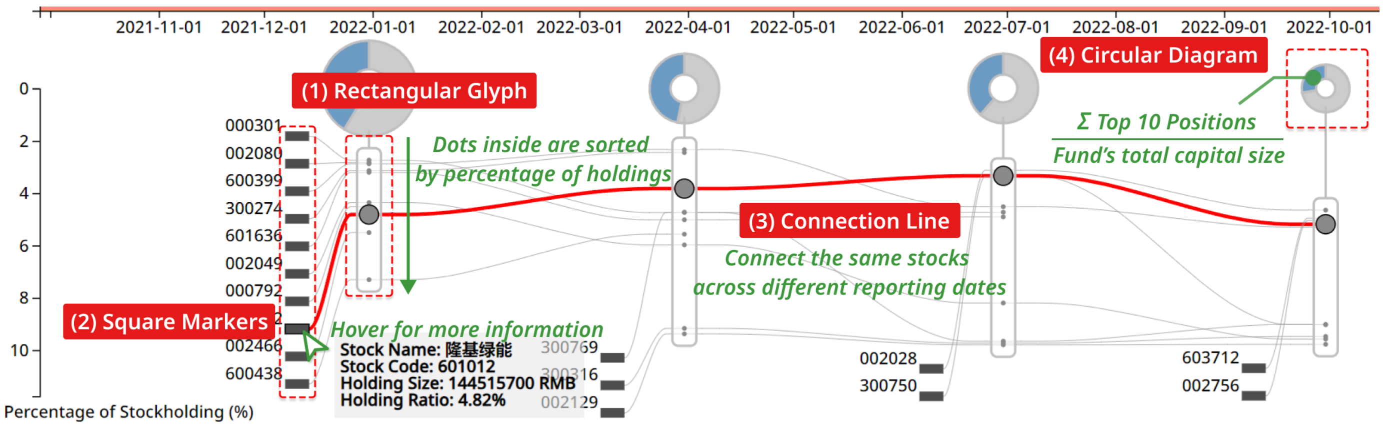


Fig. 4: Details of the *Historical Management View*: (1) The rectangular glyph is present below each reporting date. Within the glyph, each dot corresponds to a stock, representing the *Top* 10 *Positions*. The dots are arranged vertically based on the percentage of holdings, with lower dots indicating higher percentages. (2) The square marker is linked to a stock appearing for the first time for easy differentiation. Users can access more details by hovering over the marker. (3) The connection line connects the same stocks across reporting dates. (4) The radius of circular diagram corresponds to the fund's total capital size, with the blue section representing the sum of the *Top* 10 *Positions*' size as a percentage of the total capital size.

Therefore, following discussions with experts, we chose to use a table + horizon charts [55] design. As illustrated in Figure 3 (B.1), each row in the table represents a fund manager, and each column corresponds to one of the six performance metrics (*Return*, *Volatility*, *Sharpe Ratio*, *Maximum Drawdown*, *Turnover Ratio*, and *Top 10 Positions*). Above the table, a control panel is provided to empower users to customize the time range, time interval, and averaging method. The time range determines the horizontal axis range for all horizon charts in the table. The time interval governs the calculation units for the three "arbitrary interval" metrics (*Return*, *Volatility* and *Sharpe Ratio*) discussed in subsection 5.2. The averaging method adjusts the weights w_k used in computing the composite score of each fund manager, as detailed in subsection 5.2.

Semicircular Radar Chart and Ranking Control Panel. The semicircular radar chart (Figure 3 (B.2)) within the *Ranking View* features six concentric circular axes, each representing a fund manager's average performance metrics over the user-defined time range. The points, symbolizing the average of each metric for a fund manager, are connected by lines to create a polygon, where the distance from the center signifies the magnitude of the metric value. Users can engage with this chart by clicking on a row in the Metrics Table to highlight the corresponding fund manager's polygon. Furthermore, the ranking control panel (Figure 3 (B.3)) empowers users to assign weights to each metric, thereby influencing the rankings of fund managers in the metrics table. The order of rankings can be adjusted, either in ascending or descending order, by modifying the weights of the metrics. For instance, returns are typically ranked in ascending order, signifying that higher returns correspond to higher rankings.

Design Alternatives. Multiple design alternatives were considered for representing performance metrics over time, illustrated in Figure 5. The initial design employed a straightforward area chart, as shown in Figure 5 (a), where positive and negative values were differentiated by their directions, and these charts were assigned to cells in a table. To conserve space, a mirrored area chart was suggested in Figure 5 (b). However, aligning the vertical axis scales across different charts became essential when displaying performance metrics for multiple fund managers. This alignment aimed to prevent visual misinterpretation, ensuring that changes in one fund manager's performance were not visually misleading when compared with other managers' performance on the same scale. To address this challenge, a mirrored horizon graph was proposed in Figure 5 (c). This design segmented the mirrored area chart into non-overlapping bands of equal size along the vertical axis and stacked all other bands on top of the band closest to the horizontal axis. While this design resolved the issue of uniform vertical axis scales, it resulted in a flipped slope for the negative part, which was not preferred during discussions with experts. Consequently, the chosen design was an offset horizon graph, as depicted in Figure 5 (d), utilizing an offset approach rather than mirroring.

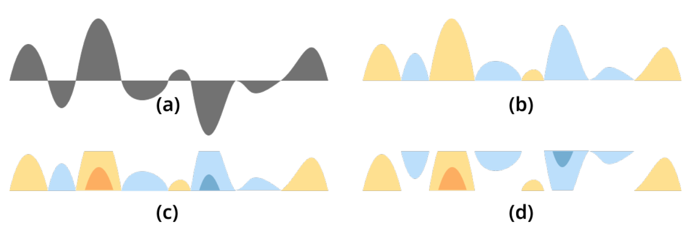


Fig. 5: Various design alternatives were explored for visualizing performance metrics over time. Initial considerations included a simple area chart (a) and a mirrored version for space efficiency (b). However, aligning vertical axis scales across different charts became crucial to prevent misinterpretation when comparing multiple fund managers. A mirrored horizon graph (c) addressed this, but its flipped slope for the negative part was not favored. Ultimately, an offset horizon graph (d) was selected, using an offset approach for a more suitable representation.

6.3 Historical Management View

The *Historical Management View* (Figure 3 (C)) provides a comprehensive review of fund managers' management records, aiding users in understanding performance metrics through fund data. When users select a fund manager in the metrics table within the *Ranking View*, all funds managed by that individual are automatically included in this view. The organized funds can be navigated vertically, and tabs at the top allow easy switching between multiple managers. Each fund is identified by its name and unique code in the upper left corner. A abbreviated timeline on the right provides a summarized view of the fund's operational history, with the red portion indicating the manager's oversight duration. Users can modify the time range on the main chart's horizontal axis (Figure 3 (C.1) & (C.2)) by brushing the abbreviated timeline.

The main chart in this view can be segmented into upper and lower subplots, both sharing a common timeline. The upper subplot (Figure 3 (C.1)) displays the fund's daily net value dynamics over time. As shown in Figure 3 (C.1), the green line signifies the fund's daily net value, and the yellow ripple area indicates the day's simulated position change compared to the previous day, with larger areas denoting more significant changes.

The lower subplot (Figure 3 (C.2)) showcases information disclosed in quarterly report. In Figure 4 (1), a rectangular glyph is present below each reporting date. Within the glyph, each dot corresponds to a stock, representing the *Top* 10 *Positions* discussed in subsection 5.2. The dots are arranged vertically based on the percentage of holdings, with lower dots indicating higher percentages. For a stock, if it becomes a member of this fund's *Top* 10 *Positions* for the first time, a square marker (Figure 4 (2)) will be linked to the stock's corresponding dot for easy differentiation. Users can access more details about the stock by hovering over the marker. Additionally, to depict the long-term holding relationship, we connect the same stocks across different reporting dates using lines (Figure 4 (3)). This feature aids users in understanding the fund's investment dynamics in a specific stock over time.

Furthermore, a circular diagram above the rectangular glyph (Figure 4 (4)) is included. The circle's radius corresponds to the fund's total capital size, with the blue section representing the sum of the *Top* 10 *Positions*' size as a percentage of the total capital size.

Design Alternatives. Figure 6 illustrates three iterations of the design aimed at presenting information in the lower subplot. In the initial version (1) depicted in Figure 6, a tree glyph design was employed. The height of the brown bar (trunk) represented the fund's size, and each circle (leaf) attached to the bar corresponded to one of the fund's *Top* 10 *Positions*. The relative location of connected circles on the bars was proportional to the percentage of holdings. However, this design did not account for the daily linkage of quarterly reports, making it challenging to track changes in the fund's holdings of specific stocks over time. To address this limitation, a second version inspired by the Sankey diagram was proposed (2) in Figure 6. Square markers below each reporting date represented the *Top* 10 *Positions*, with their vertical axis location corresponding to the percentage of holdings. Connections were drawn between the same stocks on different reporting days to facilitate tracking the dynamics of the fund's holdings. While this design addressed the initial version's problem, it faced potential visual clutter, especially when two stocks had very close investment percentages, leading to overlapping square markers. In the final design (3) (Figure 6), we opted not to use the square markers' location to encode the percentage of holdings. Instead, a layout algorithm was employed to prevent overlapping, and connections were established back to the dots in the rectangular glyphs to represent the actual percentage of holdings.

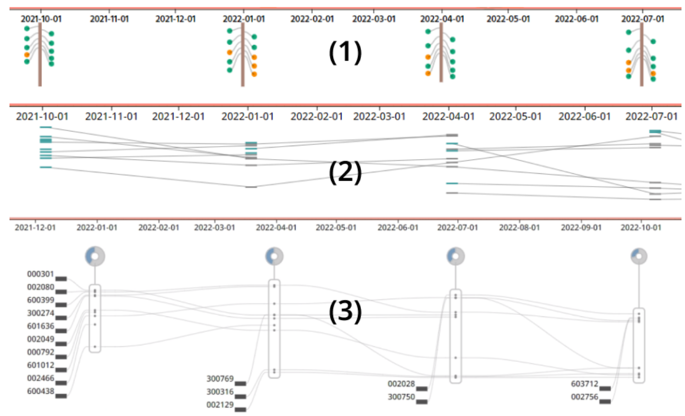


Fig. 6: Three designs were developed to present the information disclosed in quarterly reports: (1) The first version features a design where the height of the brown bar (trunk) represents the fund's size. Each circle (leaf) attached to the bar corresponds to one of the fund's *Top* 10 *Positions*. (2) The second version draws inspiration from the Sankey diagram, utilizing square markers to represent stocks. The design connects the same stocks on different reporting days, providing an effective visualization of the fund's holdings dynamics. (3) The final design combines the strengths of the first and second versions while addressing their limitations, resulting in an optimized and comprehensive representation.

6.4 Comparison View

The *Comparison View* is crafted to facilitate the comparison of fund performance among one or more fund managers, as depicted in Figure 3 (D). This view integrates a table and a line chart, and users can include funds in the *Comparison View* by selecting them within the *Historical Management View*. In the table, fundamental details about the chosen funds are presented, encompassing the fund code, manager, start management time, and end management time.

Simultaneously, the line chart visually represents the daily net value of each fund through distinct colored curves. Users can hover over funds in the table to emphasize the corresponding curves. Given that the start and end times may differ among funds, the default time range of the line chart spans from $min(all\ start\ time)$ to $max(all\ end\ time)$ to accommodate the display of all selected funds. Users retain control over the time range by brushing the abbreviated timeline below. It's important to note that if a fund's $[start\ time, end\ time]$ falls outside the brushed time range, the fund will not be visible.

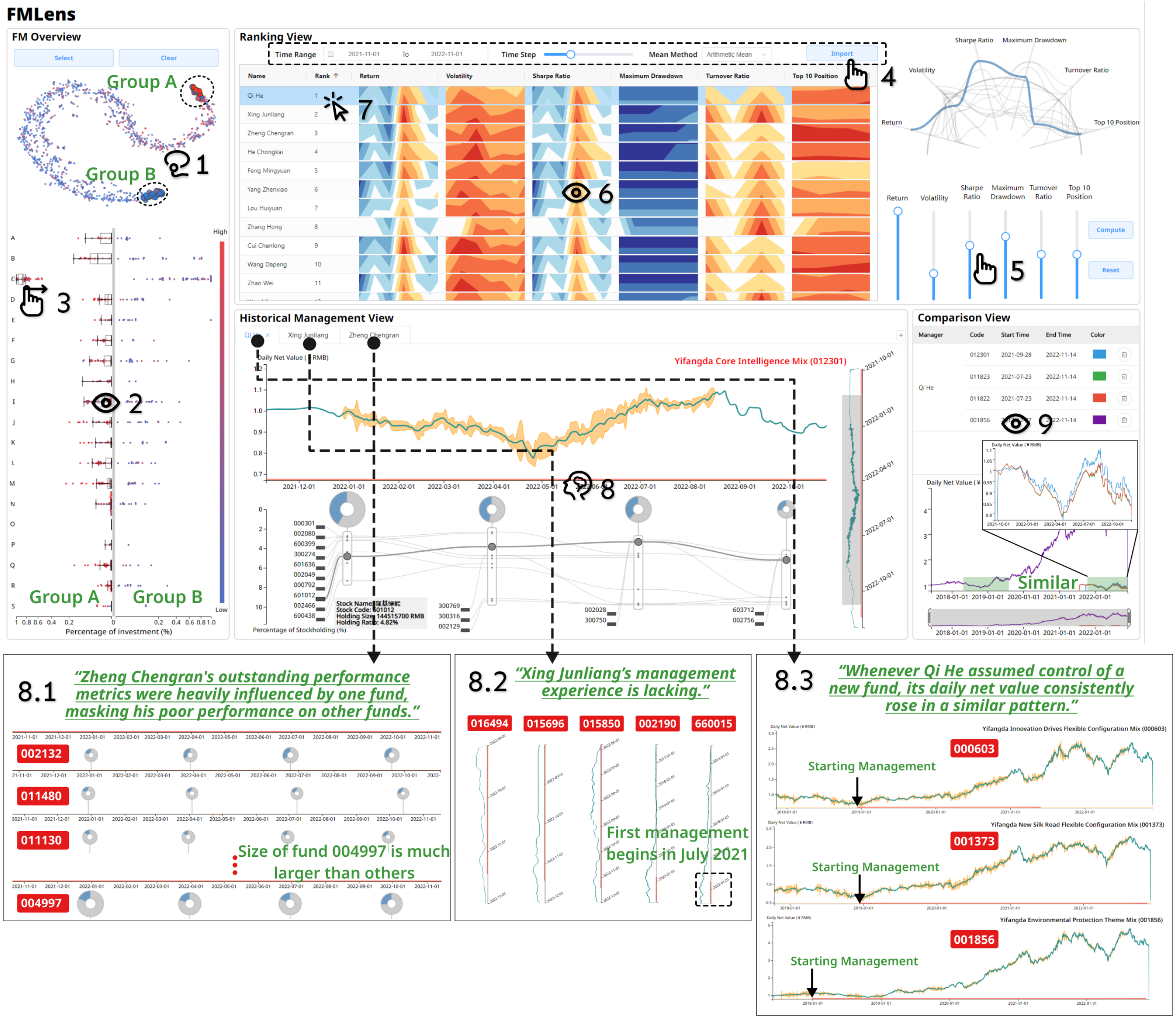


Fig. 7: Case I: (1) Lasso two groups with large differences in returns. (2) Observe the difference in investment concentration between the two groups. (3) Brush some managers for further analysis. (4) Set the time range, step and mean method to calculate metrics. (5) Adjust metric weights for ranking. (6) View the performance metrics of the top fund managers. (7) Click to check historical management records. (8) Generate findings by management records. (9) Comparison shows the funds with a similar pattern of development after the "*Qi He*" takeover.

7 Evaluation

We first showcase the efficacy of *FMLens* through two case studies, pinpointed by our collaborative experts during their system exploration. Furthermore, we conduct a user study involving 12 participants to evaluate their interaction and experience with our system.

7.1 Case Study

We invited the eight experts from subsection 3.1 who actively participated in the initial design process to experience *FMLens*. Following a brief introduction to the visual encoding and interaction logic of *FMLens*, each expert was provided with 30 − 45 minutes for unrestricted exploration of the system. The following two cases highlight some of the insights gained by the experts during their exploration.

7.1.1 Case I: A Leader in Manufacturing Investment

The first case study describes the procedure of using *FMLens* to identify a suitable fund manager for the expert E6.

Identifying return disparities. In the *FM Overview*, E6 initially observed a distinctive distribution pattern of downscaled points along a curve. Notably, she identified Group A with a concentration of red points at one end and Group B, almost entirely represented by blue points (Figure 7 (1)). Intrigued by the differences in returns between these two groups, E6 lassoed both groups for further analysis in the summary chart below. Upon closer examination, she found that Group A's investments were highly concentrated in the Manufacturing industry (industry C), while Group B exhibited a more diversified portfolio (Figure 7 (2)). E6 speculated on the reasons behind the return disparity, stating, "*manufacturing possesses the largest market size among all industry sectors, resulting in the highest potential returns during an economic upswing. Fund managers in Group A may capitalize better on the manufacturing opportunities in an economic upswing. Conversely, Group B managers may have a preference for other sectors, potentially leading to their missing out on opportunities and resulting in lower returns.*"

Analyzing top-performing fund managers. E6 then focused on fund managers in Group A with a significant percentage of investments in manufacturing. After brushing (Figure 7 (3)), these managers were added to the *Ranking View*. In this view, E6 adjusted the time range, step, and mean method for calculating performance metrics (Figure 7 (4)), and set weights for each metric based on her investment preferences (Figure 7 (5)). After generating the ranking, E6 examined the performance metrics of the top fund managers (Figure 7 (6)). She remarked, "*The results are in line with my expectations, but the performance of the top three looked very close.*" To gain deeper insights, she explored the management records of the top three in the *Historical Management View* (Figure 7 (7) & (8)).

One fund manager's strong metrics tied to one highly profitable fund. Starting with the third-ranked manager, *Zheng Chengran*, E6 discovered that while currently overseeing a substantial number of funds, only one (code: 004997) was highly profitable. Notably, the size of this fund stood out as significantly larger than any other managed by *Zheng Chengran* during the same period (Figure 7 (8.1)). E6 commented, "*Zheng Chengran's outstanding performance metrics were heavily influenced by this fund, masking his poor performance on other funds.*"

Conservative approach and limited experience contribute to another fund manager's second-place ranking. Moving on to the second-ranked *Xing Junliang*, E6 noted that this manager had only overseen five funds to date, with a track record beginning in 2021 (Figure 7 (8.2)). E6 observed that *Xing Junliang* displayed a conservative investment style with minimal position adjustments. She remarked, "*Xing Junliang's investment style is conservative, which may explain his second-place ranking. However, his management experience is lacking, and we need to monitor his future performance further.*"

The top-ranked manager consistently elevates new fund values, strenghening E6's confidence in his skills. E6 concluded her analysis by examining the history of the top-ranked manager, *Qi He*, a seasoned fund manager with almost six years of investment experience. During this exploration, E6 was surprised to find that whenever *Qi He* assumed control of a new fund, its daily net value consistently rose in a similar pattern (Figure 7 (8.3)). To validate these findings, E6 added the four funds managed by *Qi He* to the *Comparison View* (Figure 7 (9)). Remarkably, three funds managed by *Qi He* since 2021 displayed comparable development curves, closely mirroring the early trend of the exceptional fund (code: 001856) he previously managed. Ultimately, E6 asserted that *Qi He* exhibited stronger comprehensive skills and greater investment value.

7.1.2 Case II: The Present and Past Life of a Star Manager

The second case study illustrates how E4 used *FMLens* to discern the investment style of a specific fund manager.

Starting with the *FM Overview*, E4 pinpointed key data points represented by the deepest red color (indicating the highest returns) (Figure 8 (1)). This analysis revealed that a predominant preference for manufacturing existed among most fund managers. E4's interest shifted toward those deviating from this mainstream trend – those with a low percentage of manufacturing investments (Figure 8 (2)).

Uncovering one fund manager's healthcare-focused early investment style with specific stock choices. E4 proceeded to examine the overall performance of these unconventional fund managers over the past eight years. Surprisingly, this exploration led to the discovery of *Ge Lan*, an outstanding fund manager entrenched in the healthcare sector (Figure 8 (3)). Upon encountering negative news related to *Ge Lan*'s past setbacks, E4 delved into *Ge Lan*'s management records. A comprehensive examination of the net values of all funds unveiled a stark contrast in *Ge Lan*'s early performance. Notably, during her initial years, she displayed a markedly different investment style, exemplified by her third-heaviest position in *Leeco*, the addition of *Storm* to her *Top 10 Positions* in the third quarter of 2015, and the incorporation of her heaviest position at the year's end (Figure 8 (4.1)).

***Ge Lan*'s investment journey: soaring high, delisting woes, and a shift to biopharmaceutical focus.** Despite the initial market enthusiasm for stocks like *Storm* and *Leeco*, both were eventually delisted five years later. E4 pointed out that *Ge Lan* exhibited a preference for investing in high-flying companies during that period, capitalizing on valuation expansion. However, liquidity issues led to a subsequent decline in the values of these stocks. Additionally, E4 observed *Ge Lan*'s tendency to frequently change positions at the outset of her career (Figure 8 (4.1) & (4.2)), suggesting a lack of well-defined investment style. The fund's documentation, as illustrated in Figure 8 (4.2), corroborated this observation, capturing *Ge Lan*'s style transitions. Notably, E4 discerned from the stock information in Figure 8 (4.2) that *Ge Lan* eventually focused on the biopharmaceutical sector, aligning with a "long-term holdings and concentrated positions" style characteristic of the sector's dynamics.

***Ge Lan*'s steady strategy amidst growth challenges in** 2021**.** Additionally, E4 speculated on *Ge Lan*'s recent setbacks, attributing them to the imperative for fund managers to diversify as their management scale grows. Intriguingly, E4 noted that *Ge Lan* did not hastily alter her strategy in 2021 despite an increase in the size under management. As shown in Figure 8 (4.2), the persistence of positions (as indicated by the low frequency of occurrence of square markers) and the high investment percentage to *Top 10 Positions* (as indicated by the location of the dots inside the rectangular glyphs) emphasize the fact that *Ge Lan* remains entrenched in its established investment methodology in 2021 without adopting an aggressive adjustment strategy.

7.2 User Study

In order to evaluate the effectiveness of *FMLens*, a user study was conducted to compare the performance of *FMLens* with a baseline system, *Wind*.[5].

5. https://www.wind.com.cn/

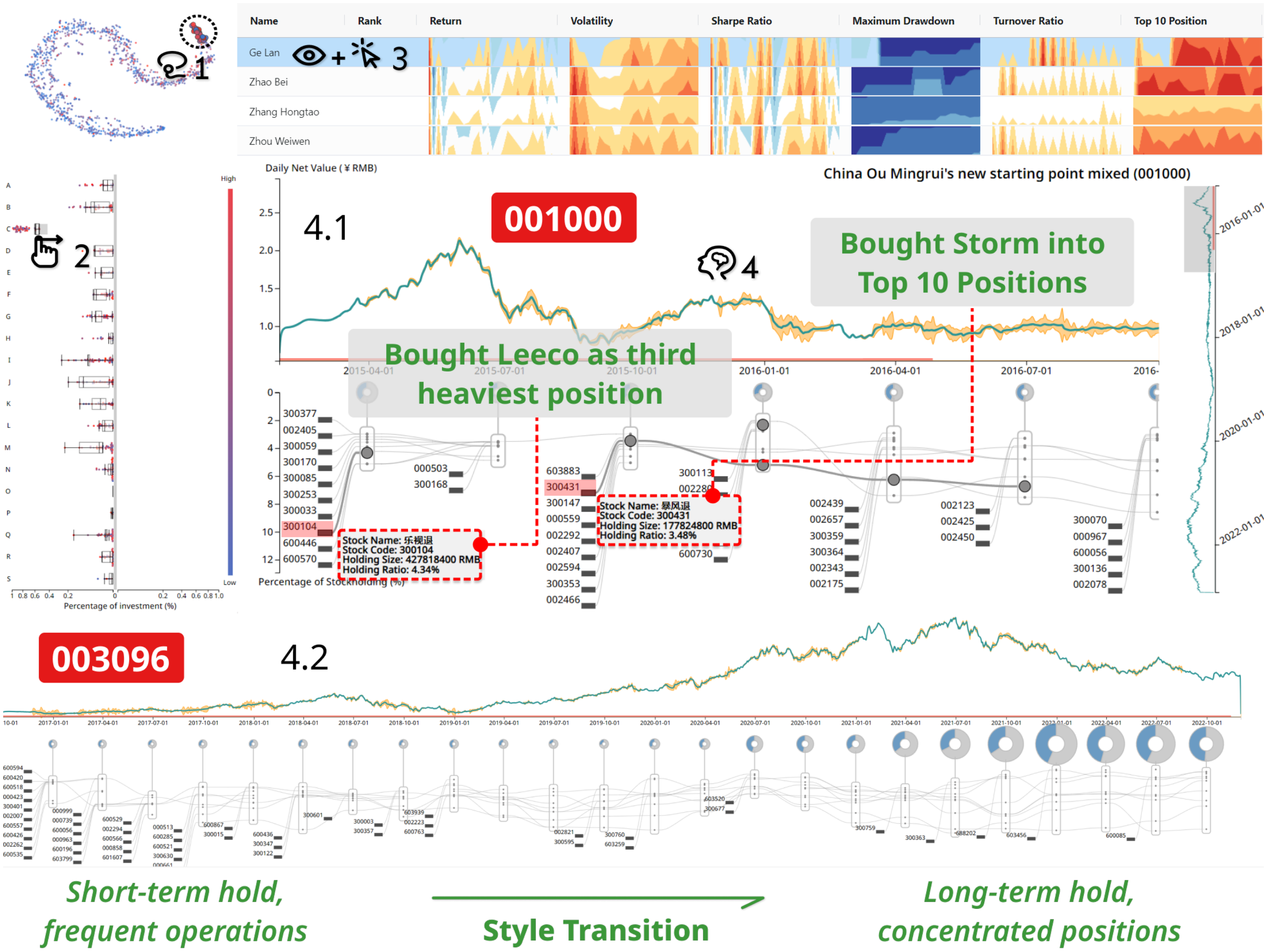


Fig. 8: Case II: (1) Select the points with the highest returns. (2) Brush fund managers with low percentage of manufacturing investments. (3) Discover an outstanding fund manager, *Ge Lan*. (4) *Ge Lan*'s early immature investment style, preferring to pursue star companies such as Leeco (4.1) and Storm (4.2). The historical records show shift in her investment style.

Participants. We recruited a total of 12 participants, each possessing a minimum of two years of experience in fund investment. None of the participants had prior exposure to our system. Among them, five individuals {P1 (Gender: male, Age: 26), P2 (male, 29), P3 (male, 33), P4 (male, 32), P5 (female, 28)} were recruited from financial institutions, where E1, E4-6 are employed, while the remaining seven participants {P6 (male, 22), P7 (female, 25), P8 (female, 24), P9 (male, 24), P10 (male, 23), P11 (male, 25), P12 (male, 27)} were recruited from universities where E2-3, E7-8 are affiliated.

Tasks and Procedure. At the beginning of the study, participants received a concise explanation of the workflows of both *FMLens* and *Wind*. This 20-minute briefing detailed the performance metrics incorporated in *FMLens*. Following the explanation, participants were provided with two links for online access to *FMLens* and *Wind*, respectively. They were then granted a 20-minute exploration period to become familiar with the features of both systems. After this exploration phase, participants engaged in the completion of four tasks, each designed to closely align with our initial design requirements.

- **T.1**: Identify a fund manager with a diversified investment sector and another with a concentrated investment sector (**R.1**).
- **T.2**: Evaluate multiple fund managers with consistently superior monthly returns throughout the 2021–2022 period (**R.2**, **R.3**).
- **T.3**: Examine the historical number of funds managed by the selected fund managers in **T.2** and analyze changes in their positions (**R.4**, **R.5**).
- **T.4**: Conduct a comparative analysis of multiple fund managers to discern variations in their investment styles (**R.5**).

Each participant was asked to complete two experiments: (1) using *FMLens* to perform **T.1**–**T.4**, and (2) using *Wind* to perform **T.1**–**T.4**. To prevent earlier experiments from influencing the results of later experiments, we randomly assigned half of the participants to complete experiment (1) before experiment (2), and the other half to complete experiment (2) before experiment (1). After completing both experiments, participants were requested to anonymously complete a 7-point Likert questionnaire, where responses spanned from 1 (strongly disagree) to 7 (strongly agree).

Feedback and comments were gathered from both sets of questionnaires. As a token of appreciation for their study participation, participants were awarded a $15 voucher.

Results. Our questionnaire aimed to evaluate and compare the performance of *FMLens* with the baseline, *Wind*, across system effectiveness, visual design convenience, and overall system usability. In the scoring results depicted in Figure 9, we employed a multiple unpaired t-test to identify significant differences between *FMLens* and *Wind*. First, the results related to system effectiveness indicate that *FMLens* surpasses *Wind* in providing an overview (Q1), checking performance metrics (Q2, Q4), adjusting metrics (Q3), personalizing rankings (Q5), tracking historical management records (Q6), understanding investment styles (Q7), and comparing returns across funds (Q8). Significant differences were observed in Q1 (p = 0.043) and Q3 (p = 0.025). P5 (female, 28 years old) highlighted, "*Although Wind provides a wealth of financial data related to fund managers, its system design lacks attention to flexible presentation and customization of information. In contrast, FMLens seems to be more convenient and efficient.*" Second, concerning visual design, participants found *FMLens* more suitable for novice users (Q9) due to its ability to provide sufficient information for evaluating fund managers (Q10). Additionally, its design and interaction were deemed helpful for the exploration and screening of fund managers (Q11). P12 (male, 27 years old) commented, "*Wind presents fund manager's performance metrics in a spreadsheet format, which is far less flexible and efficient compared to FMLens' presentation, which utilizes data visualization techniques.*" Finally, in terms of system usability, *FMLens* demonstrated greater effectiveness in identifying fund managers of interest (Q12) and comprehending the investment style and abilities of fund managers (Q13). P3 (male, 33 years old) noted, "*Wind is a data terminal lacking sufficient guidance and instructions for selecting fund managers. While FMLens' workflow is designed as a targeted tool that presents valid information step by step to help users find fund managers of interest, it is important to maintain objectivity and avoid subjective evaluations*" (Q14).

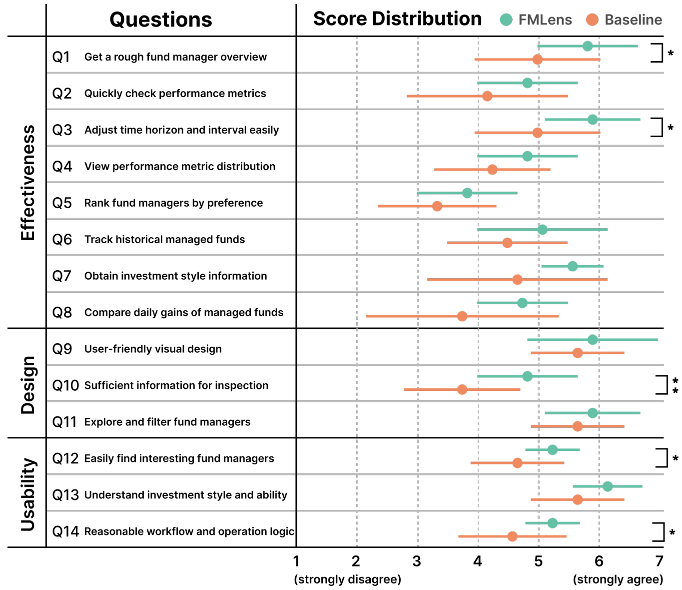


Fig. 9: The questionnaire results reflect the perspectives of 12 participants on aspects of system effectiveness, visual design convenience, and overall system usability ($* : p \leq .05, ** : p \leq .01, *** : p \leq .001, **** : p \leq .0001$).

8 Discussion, Limitation and Future Work

We conducted one-hour semi-structured interviews with E1-8 to gather their insights and suggestions regarding *FMLens*.

Contributions Over the Previous Work. This study introduces several noteworthy contributions in comparison to prior research. First, *FMLens* focuses on investing in fund managers rather than directly in funds, presenting a novel approach. As emphasized by E1, this strategy enhances the fund manager selection process, addressing a gap in support from existing tools. Second, *FMLens* offers a comprehensive evaluation of fund managers by employing a multidimensional assessment framework. Highlighted by E4, this approach integrates performance metrics and position adjustment simulations, shedding light on fund managers' typically unobservable actions. Last, the user interface of *FMLens* is deliberately crafted to be simpler and more intuitive than traditional financial data analytics platforms such as *Wind* and *Bloomberg*. Acknowledged by E7, this design feature renders the system user-friendly and easily navigable, catering to users with various investment backgrounds and experiences.

System Performance. All experts commended *FMLens* for considering multidimensional features, aiding investors in the selection of fund managers. E1 noted, "*the system is effective in specific operations, such as analyzing time-series changes in fund manager performance, which nicely compensates for the lack of existing tools.*" E2 and E3 highlighted the system's capability to assist in tracking shifts in fund managers' investment styles. E3 mentioned, "*It helps us make sense of the data in the analysis.*" E8 appreciated the design of simulated position adjustment actions, noting that it enhances understanding of the fund manager's investment philosophy and skills. Describing our system as a "*good retrospective analysis tool*," E7 stated, "*FMLens does a good job of integrating data about fund managers and presenting it in a novel way.*" Furthermore, experts mentioned that the system operates smoothly, with well-organized interactive logic.

Learning Curve and Target Users. All experts unanimously agreed that *FMLens* is user-friendly. While acknowledging that certain visual designs may require some training, experts found that an understanding of visual encoding enabled a quick initiation of exploration. They noted that the workflow, conventional yet innovative, allowed the system to support complex analytical tasks without imposing excessive barriers to use. However, E1 expressed concerns about specific performance metrics that novice investors might find challenging to comprehend. Emphasizing that even users with limited financial knowledge could evaluate fund managers based on basic metrics, E1 suggested that the system's performance could benefit from familiarity with financial matters. For example, users with minimal investment experience could easily assess fund managers based on the metric of "returns". The experts recommended that the system's target audience should comprise individuals with some investment experience rather than complete novices. Novices might find it more advantageous to select funds based on "return" rankings using existing commercial tools or apps. However, E5 also shared reservations about the system's metrics and proposed that it should include additional details about the metrics, such as calculation methods, significance, and evaluation criteria.

Generalizablility and Scalability. During the interviews, we engaged the experts in a discussion about which components of *FMLens* could be applied to other scenarios and which one(s) would require customization. They observed that *FMLens* demonstrated high versatility within the financial sector and could be effectively utilized in any financial product management organization, given appropriate pre-processing of data. E2 also suggested that there was research value in exploring the substitution of fund companies for fund managers in the system. Furthermore, experts mentioned that relevant financial data, such as stock data, could be seamlessly integrated into *FMLens*. Concerning scalability, *FMLens* efficiently presents multidimensional data on fund managers interactively. As there are currently no more than 4,000 fund managers and 10,000 funds in the entire market, and our workflow adheres to the visualization mantra of "overview first, zoom and filter, then details-on-demand" [56], there are no significant scalability issues for the front-end. The primary bottleneck is in the real-time calculation of performance metrics in the back-end. In scenarios where the user selects numerous fund managers at once or sets a small time interval, the increase in the number of calculations may lead to slower response times.

Limitation and Future Work. This study has several limitations. First, market regulations introduce a time lag in updating fund information. Although the system incorporates much of the data disclosed in quarterly reports, this information undergoes daily changes, and real-time access is unavailable. Second, there is a lack of analysis on factors influencing performance. Fund managers' performance is typically affected by various factors, including social aspects (market conditions, public sentiment, social events) and personal attributes (education, age, gender). *FMLens* does not currently support relevant attribution analysis, but we would like to consider introducing additional influencing factors in the future. Third, the current dimensionality reduction results are distributed along a curve with no clear clustering pattern. In order to improve the effectiveness of clustering post-dimensionality reduction, additional attributes such as education level and years of employment can be considered in the future to summarize the candidate space. Fourth, users might not be attuned to custom weights. In fund manager ranking, users are required to customize weights for different performance metric attributes. For some users, this may not be a familiar ranking approach, and guidance is essential to understand and optimize weight settings, supporting a more reliable and contextual decision-making process. Fifth, the fund position simulation method relies solely on daily fund return data and industry index data. In the future, we would like to try to introduce more financial data for position simulation. Additionally, the experiment design in this study utilizes a small sample size (90 equity funds) to demonstrate the methodology's validity, which may not ensure the same accuracy with a larger sample size. In future work, we would like to perform experiments on larger samples to verify the accuracy of the method. Sixth, the evaluation focuses on qualitative feedback and lacks quantitative analysis to validate the impact of the system on improving fund manager selection outcomes. This is because fund manager performance is often influenced by a variety of long-term factors. To effectively assess their performance, a comprehensive and long-term tracking mechanism needs to be established, which requires sustained resource investment.

9 Conclusion

This study introduces a visual analytics approach designed to enhance the efficiency of the fund manager selection process and empower investors to assess potential fund managers more effectively. By leveraging *FMLens*, investors can perform a comparative analysis of investment styles, abilities, and performance metrics across various fund managers, dynamically ranking them based on individual preferences. The efficacy of *FMLens* is validated through two case studies and a qualitative user study.

Acknowledgments

We would like to express our gratitude to our domain experts and the anonymous reviewers for their insightful comments. This work is funded by grants from the National Natural Science Foundation of China (No. 62372298), the Shanghai Frontiers Science Center of Human-centered Artificial Intelligence (ShangHAI), and the Key Laboratory of Intelligent Perception and Human-Machine Collaboration (ShanghaiTech University), Ministry of Education.

References

[1] R. C. Jones and R. Wermers, "Active management in mostly efficient markets," *Financial Analysts Journal*, vol. 67, no. 6, pp. 29–45, 2011.

[2] K. P. Baks, *On the performance of mutual fund managers*. University of Pennsylvania, 2002.

[3] J. B. Berk and J. H. Van Binsbergen, "Measuring skill in the mutual fund industry," *Journal of financial economics*, vol. 118, no. 1, pp. 1–20, 2015.

[4] A. A. Gottesman and M. R. Morey, "Manager education and mutual fund performance," *Journal of empirical finance*, vol. 13, no. 2, pp. 145–182, 2006.

[5] M. Grinblatt and S. Titman, "Performance measurement without benchmarks: An examination of mutual fund returns," *Journal of business*, pp. 47–68, 1993.

[6] W. F. Sharpe, "Asset allocation: Management style and performance measurement," *Journal of portfolio Management*, vol. 18, no. 2, pp. 7–19, 1992.

[7] J. A. Christopherson, W. E. Ferson, and D. A. Glassman, "Conditioning manager alphas on economic information: Another look at the persistence of performance," *The Review of Financial Studies*, vol. 11, no. 1, pp. 111–142, 1998.

[8] J. Chevalier and G. Ellison, "Are some mutual fund managers better than others? cross-sectional patterns in behavior and performance," *The journal of finance*, vol. 54, no. 3, pp. 875–899, 1999.

[9] E. F. Fama and K. R. French, "Luck versus skill in the cross-section of mutual fund returns," *The journal of finance*, vol. 65, no. 5, pp. 1915–1947, 2010.

[10] M. Kacperczyk, C. Sialm, and L. Zheng, "Unobserved actions of mutual funds," *The Review of Financial Studies*, vol. 21, no. 6, pp. 2379–2416, 2008.

[11] W. F. Sharpe, "Mutual fund performance," *The Journal of business*, vol. 39, no. 1, pp. 119–138, 1966.

[12] M. C. Jensen, "The performance of mutual funds in the period 1945-1964," *The Journal of finance*, vol. 23, no. 2, pp. 389–416, 1968.

[13] F. A. Sortino and L. N. Price, "Performance measurement in a downside risk framework," *the Journal of Investing*, vol. 3, no. 3, pp. 59–64, 1994.

[14] A. Niessen-Ruenzi and S. Ruenzi, "Sex matters: Gender bias in the mutual fund industry," *Management Science*, vol. 65, no. 7, pp. 3001–3025, 2019.

[15] K. M. Cremers and A. Petajisto, "How active is your fund manager? a new measure that predicts performance," *The review of financial studies*, vol. 22, no. 9, pp. 3329–3365, 2009.

[16] R. Chang, M. Ghoniem, R. Kosara, W. Ribarsky, J. Yang, E. Suma, C. Ziemkiewicz, D. Kern, and A. Sudjianto, "Wirevis: Visualization of categorical, time-varying data from financial transactions," in *2007 IEEE Symposium on Visual Analytics Science and Technology*. IEEE, 2007, pp. 155–162.

[17] A. Samak, H. C. Lam, B. D. Fisher, and D. S. Ebert, "An experimental study of financial portfolio selection with visual analytics for decision support," *2011 44th Hawaii International Conference on System Sciences*, pp. 1–10, 2011.

[18] S. Ko, I. Cho, S. Afzal, C. Yau, J. Chae, A. Malik, K. Beck, Y. Jang, W. Ribarsky, and D. S. Ebert, "A survey on visual analysis approaches for financial data," in *Computer Graphics Forum*, vol. 35, no. 3. Wiley Online Library, 2016, pp. 599–617.

[19] E. V. F. Lapura, J. K. J. Fernandez, M. J. K. Pagatpat, and D. D. Dinawanao, "Development of a university financial data warehouse and its visualization tool," *Procedia Computer Science*, vol. 135, pp. 587–595, 2018.

[20] A. Arleo, C. Tsigkanos, C. Jia, R. A. Leite, I. Murturi, M. Klaffenböck, S. Dustdar, M. Wimmer, S. Miksch, and J. Sorger, "Sabrina: Modeling and visualization of financial data over time with incremental domain knowledge," in *2019 IEEE Visualization Conference (VIS)*. IEEE, 2019, pp. 51–55.

[21] M. Schaefer, F. Wanner, R. Kahl, L. Zhang, T. Schreck, and D. Keim, "A novel explorative visualization tool for financial time series data analysis," in *VAW2 011: The Third International UKVAC Workshop on Visual Analytics*, 2011.

[22] H. Ziegler, M. Jenny, T. Gruse, and D. A. Keim, "Visual market sector analysis for financial time series data," in *2010 IEEE Symposium on Visual Analytics Science and Technology*. IEEE, 2010, pp. 83–90.

[23] E. Sorenson and R. Brath, "Financial visualization case study: Correlating financial timeseries and discrete events to support investment decisions," in *2013 17th International Conference on Information Visualisation*. IEEE, 2013, pp. 232–238.

[24] A. Jungmeister, "Adapting treemaps to stock portfolio visualization," *UMD HCIL*, 03 1991.

[25] C. Csallner, M. Handte, O. Lehmann, and J. Stasko, "Fundexplorer: Supporting the diversification of mutual fund portfolios using context treemaps," in *IEEE Symposium on Information Visualization 2003 (IEEE Cat. No. 03TH8714)*. IEEE, 2003, pp. 203–208.

[26] X. Yue, J. Bai, Q. Liu, Y. Tang, A. Puri, K. Li, and H. Qu, "sportfolio: Stratified visual analysis of stock portfolios," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 1, pp. 601–610, 2019.

[27] X. Yue, Q. Gu, D. Wang, H. Qu, and Y. Wang, "iquant: Interactive quantitative investment using sparse regression factors," *arXiv preprint arXiv:2104.11485*, 2021.

[28] H. Guo, M. Liu, B. Yang, Y. Sun, H. Qu, and L. Shi, "Rankfirst: Visual analysis for factor investment by ranking stock timeseries," *IEEE Transactions on Visualization and Computer Graphics*, 2022.

[29] T. Dwyer, "A scalable method for visualising changes in portfolio data," in *Proceedings of the Asia-Pacific symposium on Information visualisation-Volume 24*, 2003, pp. 17–25.

[30] J. Li, H. Izakian, W. Pedrycz, and I. Jamal, "Clustering-based anomaly detection in multivariate time series data," *Applied Soft Computing*, vol. 100, p. 106919, 2021.

[31] A. Abdella, J. K. Brecht, and I. Uysal, "Statistical and temporal analysis of a novel multivariate time series data for food engineering," *Journal of Food Engineering*, vol. 298, p. 110477, 2021.

[32] B. D. Nguyen, R. Hewett, and T. Dang, "Netscatter: Visual analytics of multivariate time series with a hybrid of dynamic and static variable relationships," in *2021 IEEE 14th Pacific Visualization Symposium (PacificVis)*. IEEE, 2021, pp. 52–60.

[33] S. T. Lei and K. Zhang, "Visual signatures for financial time series," in *Proceedings of the 2011 Visual Information Communication-International Symposium*, 2011, pp. 1–10.

[34] H. Hochheiser and B. Shneiderman, "Dynamic query tools for time series data sets: timebox widgets for interactive exploration," *Information Visualization*, vol. 3, no. 1, pp. 1–18, 2004.

[35] Y. Wu, N. Pitipornvivat, J. Zhao, S. Yang, G. Huang, and H. Qu, "egoslider: Visual analysis of egocentric network evolution," *IEEE transactions on visualization and computer graphics*, vol. 22, no. 1, pp. 260–269, 2015.

[36] R. Peng, "A method for visualizing multivariate time series data," *Journal of Statistical Software, Code Snippets*, vol. 25, no. 1, p. 1–17, 2008.

[37] Y. Chen, Q. Chen, M. Zhao, S. Boyer, K. Veeramachaneni, and H. Qu, "Dropoutseer: Visualizing learning patterns in massive open online courses for dropout reasoning and prediction," in *2016 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 2016, pp. 111–120.

[38] Q. Li, X. Wei, H. Lin, Y. Liu, T. Chen, and X. Ma, "Inspecting the running process of horizontal federated learning via visual analytics," *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[39] D. J. Hruschka, D. Schwartz, D. C. St. John, E. Picone-Decaro, R. A. Jenkins, and J. W. Carey, "Reliability in coding open-ended data: Lessons learned from hiv behavioral research," *Field methods*, vol. 16, no. 3, pp. 307–331, 2004.

[40] W. F. Sharpe, "Capital asset prices: A theory of market equilibrium under conditions of risk," *The journal of finance*, vol. 19, no. 3, pp. 425–442, 1964.

[41] Y.-J. Lai, T.-Y. Liu, and C.-L. Hwang, "Topsis for modm," *European journal of operational research*, vol. 76, no. 3, pp. 486–500, 1994.

[42] R. D. Harris and M. Mazibas, "Dynamic hedge fund portfolio construction: A semi-parametric approach," *Journal of Banking & Finance*, vol. 37, no. 1, pp. 139–149, 2013.

[43] B. Murthi, Y. K. Choi, and P. Desai, "Efficiency of mutual funds and portfolio performance measurement: A non-parametric approach," *European Journal of Operational Research*, vol. 98, no. 2, pp. 408–418, 1997.

[44] M. M. L. de Prado, *Machine learning for asset managers*. Cambridge University Press, 2020.

[45] J. Huang, J. Chai, and S. Cho, "Deep learning in finance and banking: A literature review and classification," *Frontiers of Business Research in China*, vol. 14, no. 1, pp. 1–24, 2020.

[46] J. B. Heaton, N. G. Polson, and J. H. Witte, "Deep learning for finance: deep portfolios," *Applied Stochastic Models in Business and Industry*, vol. 33, no. 1, pp. 3–12, 2017.

[47] M. Arouri, H. Ben Ameur, N. Jawadi, F. Jawadi, and W. Louhichi, "Are islamic finance innovations enough for investors to escape from a financial downturn? further evidence from portfolio simulations," *Applied Economics*, vol. 45, no. 24, pp. 3412–3420, 2013.

[48] R. Fernholz, R. Garvy, and J. Hannon, "Diversity-weighted indexing," *Journal of Portfolio Management*, vol. 24, no. 2, p. 74, 1998.

[49] R. Liu, J. Kuang, Q. Gong, and X. Hou, "Principal component regression analysis with spss," *Computer methods and programs in biomedicine*, vol. 71, no. 2, pp. 141–147, 2003.

[50] G. C. McDonald, "Ridge regression," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 1, no. 1, pp. 93–100, 2009.

[51] J. Ranstam and J. Cook, "Lasso regression," *Journal of British Surgery*, vol. 105, no. 10, pp. 1348–1348, 2018.

[52] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance," *Climate research*, vol. 30, no. 1, pp. 79–82, 2005.

[53] F. Heimerl and M. Gleicher, "Interactive analysis of word vector embeddings," in *Computer Graphics Forum*, vol. 37, no. 3. Wiley Online Library, 2018, pp. 253–265.

[54] P. E. Rauber, S. G. Fadel, A. X. Falcao, and A. C. Telea, "Visualizing the hidden activity of artificial neural networks," *IEEE transactions on visualization and computer graphics*, vol. 23, no. 1, pp. 101–110, 2016.

[55] H. Reijner *et al.*, "The development of the horizon graph," in *Proc. Vis08 Workshop From Theory to Practice: Design, Vision and Visualization*, vol. 3, 2008.

[56] B. Shneiderman, "The eyes have it: A task by data type taxonomy for information visualizations - sciencedirect," *The Craft of Information Visualization*, pp. 364–371, 2003.

Longfei Chen is currently in the third year of his M.S. program in Computer Science and Technology at ShanghaiTech University. He actively contributes as a member of ViSeer LAB under the mentorship of Prof. Quan Li. Longfei's academic interests encompass urban data visualization, eXplainable Artificial Intelligence (XAI), and Financial Technology (FinTech). His research primarily focuses on enhancing the understanding and communication of data and models by employing scalable visual interfaces and interpretable methodologies. For additional information, please visit https://chenlf126.github.io/.

Chen Cheng is a senior undergraduate student at the School of Information Science and Technology at ShanghaiTech University, with a specialization in Human-AI Interaction (HAI). In the past, he has been involved in the design and construction of interactive systems for visualization and analysis, working under the guidance of Prof. Quan Li at ShanghaiTech University. These systems are designed to enhance the understanding and decision-making processes when dealing with extensive datasets. Additionally, Chen Cheng has contributed to the development of a programming support tool for code comprehension, collaborating with Prof. Tianyi Zhang at Purdue University.

He Wang is currently enrolled in the M.S. program at the School of Information Science and Technology, ShanghaiTech University. He holds a B.S. degree from the School of Computer Science and Information Engineering at Hefei University of Technology. His research interests encompass data visualization, data-centric AI, and eXplainable Artificial Intelligence (XAI). His overarching research objective is to advance human-AI collaboration by creating scalable and interpretable visual analytics tools.

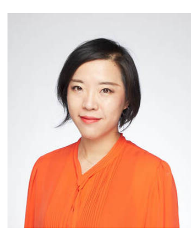
Xiyuan Wang is presently pursuing a Master's degree in Computer Science at ShanghaiTech University under the supervision of Prof. Quan Li. She holds a B.S. degree from the same university. Xiyuan's research focuses on human-computer interaction (HCI) and large data visualization (VIS). Additionally, she is actively involved in exploring research areas related to storytelling.

Yun Tian currently serves as a Research and Development Engineer with expertise in data visualization at CFETS Information Technology (Shanghai) Co., Ltd, a financial technology company. He earned his master's degree from the School of Information Science and Technology at ShanghaiTech University. Yun Tian holds a B.S. degree from the School of Astronomy and Space Science at Nanjing University. His professional focus includes heterogeneous networks, graph neural networks, and visual analytics.

Xuanwu Yue currently holds the position of Senior Data Analytics at Nanjing AI Institute of Sinovation Ventures. Additionally, he is affiliated with Baiont Capital, a leading AI hedge fund primarily focusing on Chinese markets. His primary research area involves the visual analytics of financial data, with specific emphasis on quantitative trading, market analysis, and high-frequency trading.

Wong Kam-Kwai is currently pursuing a Ph.D. in the Department of Computer Science and Engineering at the Hong Kong University of Science and Technology (HKUST). He obtained his B.E. degree from HKUST. His primary research interests encompass data visualization, visual analytics, and data mining. For further information, please visit https://kamkwai.com.

Haipeng Zhang serves as an assistant professor in the School of Information Science and Technology at ShanghaiTech University. His research centers on mining and modeling large-scale human behavioral data, with a particular emphasis on investigating patterns related to human investment, communication, mobility, and innovation. These studies aim to provide insights and implications across financial, political-economic, and scientific domains. Haipeng Zhang has published in prestigious venues such as WWW, WSDM, TKDD, CIKM, ICLR, and AAAI, and his work has garnered media coverage from notable outlets like the New Scientist magazine and the Communications of the ACM website.



Suting Hong serves as an assistant professor at the School of Entrepreneurship and Management at ShanghaiTech University. Her research focuses on Entrepreneurial Finance and Start-up Innovations. Utilizing both theoretical models and empirical analysis, she explores various factors influencing the development of tech-oriented start-ups.

Quan Li servers as a tenure-track assistant professor position at the School of Information Science and Technology, ShanghaiTech University. He earned his Ph.D. from the Department of Computer Science and Engineering at the Hong Kong University of Science and Technology. His research spans artificial intelligence, big data visualization, visual analysis, interpretable machine learning, data science, data storytelling, human-computer interaction (HCI), and computational social sciences. Further information can be found at https://faculty.sist.shanghaitech.edu.cn/liquan/.