



A decision support system for stock investment recommendations using collective wisdom

Jörg Gottschlich^{*}, Oliver Hinz¹

Technische Universität Darmstadt, Hochschulstraße 1, 64289 Darmstadt, Germany



ARTICLE INFO

Article history:

Received 7 February 2013

Received in revised form 2 September 2013

Accepted 18 October 2013

Available online 25 October 2013

Keywords:

Wisdom of crowds

Investment decision support system

Virtual investment communities

Portfolio creation

ABSTRACT

Previous research has shown that user-generated stock votes from online communities can be valuable for investment decisions. However, to support investors on a day-to-day basis, there is a need for an efficient support system to facilitate the use of the data and to transform crowd votes into actionable investment opportunities. We propose a decision support system (DSS) design that enables investors to include the crowd's recommendations in their investment decisions and use it to manage a portfolio. A prototype with two test scenarios shows the potential of the system as the portfolios recommended by the system clearly outperform the market benchmark and comparable public funds in the observation period in terms of absolute returns and with respect to the Reward-to-Variability-Ratio.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction and motivation

The rise of user-generated content on the Internet enabled a wider public to participate in online content creation and publication without the need for deep technical expertise. The technical possibility to centrally aggregate the local contributions of a large crowd enables the creation of artifacts which are of equal or superior quality than those made by experts in the domain. Wikipedia, as an example, reaches a comparable quality to the renowned Britannica [19], solely depending on the contributions of a diverse anonymous crowd. This effect, coined as “wisdom of crowds” [41], is based on the diversity in information possession and processing of the individual members and is evident in a number of problem solving situations such as judging, estimating or decision making [18,28].

Estimation tasks are a problem class where group judgments prove to perform extraordinarily well. The reason behind is an effect called *bracketing* [39], which refers to the high likelihood that a part of the crowd will overestimate, while another part will underestimate the true value. Hence, averaging all judgments will lead to a more accurate judgment than that of the average judge [26]. Take an example: two people estimate the outside temperature for the next day as 60 °F and 80 °F, while the true temperature will be 73 °F. The estimates were wrong by 13 °F and 7 °F, or 10 °F on average. However, the mean of the two estimates, 60 °F and 80 °F, which is 70 °F, is off by only 3 °F. So

using deviation as a measure, the average *judgment* outperforms the average *judge* [cf. 26].

One form of harnessing the crowd wisdom to improve decision making is the application of prediction markets. At prediction markets, participants can buy or sell contracts whose payoff is connected to a certain future event, e.g. “Candidate A will win the election”. By dealing contracts over time, the contracts' prices reflect the market participants' collective judgment of the likelihood that the associated event will become true. The collective judgment has been proven to be quite close to the final result [3,18,40]. If designed appropriately, such prediction markets can be utilized to support decisions [7]. Preference markets are a closely related concept and have been used to apply the wisdom of crowd to evaluate emerging technologies at an early stage of product development [10,15]. The difficulty here is to prioritize resources for technologies who are most promising and which might emerge into product features. This problem is an instance of a typical investment problem and hence very close to the task in focus of this paper, the beneficial allocation of capital to capital market shares. Chen et al. [10] used a preference market to compare the crowd estimate of product feature ranking to a benchmark ranking done by an expert group and found indications that they reach comparable results if the preference market provides a sufficient (non-monetary) incentive. In the perspective of the financial domain, this finding suggests that trading strategies based on crowd recommendations might be able to perform as well as public funds managed by experienced domain experts even without a direct monetary reward for the crowd members. Specifically, virtual investing communities (VICs) (see Section 1.1) which collect user opinions on stock development usually provide non-monetary incentives for participation such as public reputation or access to exclusive information.

^{*} Corresponding author. Tel.: +49 6151 1675223; fax: +49 6151 1672220.

E-mail addresses: gottschlich@emarkets.tu-darmstadt.de (J. Gottschlich), hinz@emarkets.tu-darmstadt.de (O. Hinz).

¹ Tel.: +49 6151 1675220; fax: +49 6151 1672220.

ACI WORLDWIDE INC.

US0044981019 | 02/04/2013



Fig. 1. Example of VIC vote.

1.1. Wisdom of crowd in finance

With respect to the financial domain, there is extensive research on the value of user-generated content for stock investment decisions. Antweiler and Frank [2] analyze the information content of stock discussion boards and find evidence that message posts can be used to predict stock market trading volume and volatility and – to a small extent – stock returns. Mood analysis from Twitter messages can be used to improve prediction accuracy of the Dow Jones Industrial Average Index [8].

Stock discussion boards evolved into Internet portals, so called virtual investing communities (VICs), where members are able to provide their guess about a share's future performance in a structured way (see Fig. 1 for an example). Members can make a buy or sell recommendation for any share along with a target price at a specific future date (= vote). By doing so, the individual investors act like institutional analysts and the aggregation of the single votes of a share leads to a collective judgment of its prospects. Examples for such websites are CAPS² or sharewise.³ As participants make very specific (price) predictions of a share, those platforms can be seen as a prediction market for share prices and are as such an enhancement of stock discussion boards with unstructured free-text information [4] that has to be preprocessed for analysis (including a certain loss of accuracy).

There is evidence in literature that information from these stock communities can be used to implement profitable stock investment strategies. Using data of the CAPS platform, Avery et al. (2011) find that stocks ranked highest by the community indeed show a better subsequent performance than those that were ranked low [4]. Especially short (i.e. sell) recommendations of the crowd are able to predict stock price declines. Further, they analyze the composition of the performance and find that the advantage of the crowd comes from stock selection rather than market timing or style/risk factors (as identified by Fama and French [17] and Carhart [9] which classify stocks by their market or risk profile).

An in-depth analysis of the CAPS data for investment purposes can be found at Hill and Ready-Campbell [22]. They find evidence, too, that a portfolio based on crowd voting is able to outperform the market index (S&P 500) and that the higher rated shares do indeed perform better. Specifically, they find that a crowd of about 250 people always outperforms the S&P 500 index. In addition, they rank the users according to their past performance and find that a selected group of experts from the crowd performs better than the whole crowd. They test several investment strategies for portfolio construction, but disregard transaction cost. While their analysis is comprehensive and their results are insightful, investors who want to make use of the effects for future scenarios are left with extensive analytical effort and not much guidance on how to transform results into actionable investment decisions.

Nofer and Hinz (2013) empirically show that the average institutional expert from the financial service industry and the average private

crowd member are able to outperform the market. More surprisingly they also show that investors are on average significantly better off when trusting a crowd recommendation than following the advice of professional experts from banks. However, the authors do not provide and evaluate a system that implements their findings and which offers decision support for investors [34].

Making use of the value in crowd data for day-to-day investment decision is not a trivial task. Because analyses are complex (both in terms of method and data) and time-consuming when done manually, a decision support system proves to be helpful.

1.2. DSS for investment decisions

There is a vast amount of literature about systems designed to support stock investment decisions with a large diversity in focus and approach. One stream focuses on asset and liability management (ALM) topics suited for professional institutions like banks which seek support for risk management comprising all of their asset classes. Moynihan et al. (2002) suggest a DSS that forecasts the amount of assets and liabilities and the primary interest rate and involve simulation models to conduct gap analysis and rate risk of the institution. Additionally, it is possible to run “what-if” scenarios to analyze developments under changing market conditions [32]. A more recent approach utilizes complex stochastic programming methods to support optimal strategic asset allocation providing a user-friendly web interface [6].

In regard to the topic of stock investment, the majority of previous research strives to provide better insights to investors by improved information support. Methods to model stock price development using optimization or machine-learning approaches are commonly used. Specifically, artificial neural networks show broad coverage in investment decision support literature, often in combination with other approaches. Tsaih et al. (1998) combine a neural network approach with a static rule base to predict the direction of daily price changes in the S&P 500 stock index futures which outperforms a passive buy-and-hold strategy [43]. Chou et al. (1996) follow a similar approach for the Taiwanese market [12]. Liu and Lee (1997) propose an Excel-based tool for technical analysis [27]. Other approaches combine neural networks with genetic algorithms [e.g. Ref. 5]. Kuo et al. (2001) show that they reach a higher prediction accuracy of stock development when including qualitative factors (e.g. political effect) in addition to quantitative data [25].

More interactive forms of decision support, where systems provide a laboratory-like environment for prospective investors to conduct standard as well as customized analyses, have appeared. Dong et al. (2004) suggest a framework for a web based DSS which implements a comprehensive approach to the investment task up to rebalancing an investor's portfolio according to his risk/return profile. They integrated On-Line Analytical Processing (OLAP) tools for customized multidimensional analyses [16]. Another approach for interactive investor support provides the possibility of stepwise model generation such that investors can start with simple models from a toolbox and incrementally add more building blocks to arrive at more complex prediction models

² <http://caps.fool.com/>.

³ <http://www.sharewise.com/>.

[11]. With a focus on more actionable support for private investors, Muntermann suggests a system design that estimates price effects of ad hoc notifications for public companies and sends out text messages to mobile phones including predicted effect size and time window [33].

1.3. Purpose of this paper

In summary, there is a broad range of investment decision support systems, ranging from providing better informational insight for the investor to very specific actionable investment support. To our knowledge, no investment decision support system uses crowd votes from online stock community platforms to derive investment decisions even though research indicates that valuable insights can be expected. Therefore, we develop a system that is capable of supporting investors in their daily task of selecting promising stocks for their portfolio. Furthermore, the system is able to select stocks based on the crowd votes by an individual investment strategy and to transform those picks into a target portfolio. We provide first application results that demonstrate the advantages for investors to make use of such a DSS.

The remainder of this article is organized as follows: We present the decision support process and provide the system design that implements this process in Section 2. The implementation of the design as a prototype can be found in Section 3 and the evaluation of two test scenarios in Section 4, while we summarize our contribution in Section 5.

2. System design

2.1. Methodology

In this paper we follow the methodology of Design Science Research as suggested by Hevner et al. (2004) which confounds an approach rooted in the engineering sciences. They provide several guidelines which we follow in the development of the proposed architecture [21]:

- Design as an artifact: Design-science research must produce a viable artifact in the form of a construct, a model, a method, or an instantiation.
- Problem relevance: The objective of design-science research is to develop technology-based solutions to important and relevant business problems.
- Design evaluation: The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods.
- Research contributions: Effective design-science research must provide clear and verifiable contributions in the areas of the design artifact, design foundations, and/or design methodologies.
- Research rigor: Design-science research relies upon the application of rigorous methods in both the construction and evaluation of the design artifact.
- Design as a search process: The search for an effective artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment.
- Communication of research: Design-science research must be presented effectively both to technology-oriented as well as management-oriented audiences.

The introduction and the theoretical foundation in Section 1 show the relevance of the topic in the context of research. The relevance of the business problem is obvious, as outperforming the market is both relevant to the whole financial industry and valuable given the profit prospects. Our artifact, the decision support system, addresses this problem and provides a solution (Section 2) which is the result of the rigor search based on theory for a solution to the identified problem of how to make use of user-generated content for investment decision support. We evaluate our design and show its functionality in Sections 3 and 4. In our concluding remark (Section 5) we summarize our research

contribution. All in all, this paper communicates our research to foster discussion as well as further research and applications in this area.

2.2. Crowd stock voting data structure

We used a data set of recommendations for a certain ISIN (International Securities Identification Number) on a certain day (called “votes”) from an online stock community site. The main components of such a vote are: Stock Identification (i.e. ISIN), Recommendation (Buy/Sell), Start date (publication date of vote), End date, Start price (price of stock at time of vote publication), Target price (forecasted price), and End price (actual price at close of vote). The time from start date to end date is the runtime of a vote and it is called “open” on any date in between. Within this time frame, the target price is expected to be met. If the member who creates a vote does not specify an end date, the vote is automatically closed by the platform after 180 days. In addition, the member can close a vote at any time before the end date is reached. Recommendations can either be *buy* or *sell*, the target price accordingly is above or below the price at the publication of the vote (= start price). Fig. 1 shows an example of such a user vote as it appears on the site.

2.3. System design and decision support process

The system can support the investor in three different aspects: First, by creating a ranked list, the investor gets advice about the most preferable securities on a specific date. This information alone is already useful when selecting securities for a portfolio. In addition, the system supports the implementation and simulation of strategies based on the computed ranking so that investors can test and explore different approaches to identify promising investment strategies. Once a suitable strategy has been identified, the system can be used to automatically follow a specified strategy day by day and create orders to modify a portfolio. This functionality helps to stick to a rational strategy and prevent investor's emotions to interfere and bias a well-defined strategy [cf. 29].

The system's task is to transform crowd votes into actionable share ratings for a given day. We split this task in two phases: 1) Rating computation: Translate crowd votes into one or several performance estimate metrics per ISIN and 2) Investment phase: Building a portfolio based on a list of shares ranked by such a metric, from which we select shares to create a portfolio. This approach follows the pattern of Ref. [22]. Fig. 2 gives an overview of the system design. The system executes both rating computation and investment phase for each day in the focus period.

2.3.1. Rating computation

The rating computation builds upon three data sources: the crowd votes, share prices for the period in focus and additional share metadata such as industry or geographic information. The system accesses these data either from a local database or online (e.g. via calls to APIs⁴ of data providers). The term *rating*, as we use it, refers to the set of all computed metrics for a certain ISIN on a certain day. Which components of a rating are taken into account for portfolio creation depends on the investment strategy. The available metrics form a library of metrics as a basis for investment strategies.

Our design offers two approaches to arrive at a performance rating from the crowd votes. Simple metrics are computed by the system itself, without the need of external statistical software. By that, we avoid the overhead and effort of involving complex software packages when the metrics only involve simple average or sum calculations. However, in addition, we attach a modeling facility that supports the application of machine learning techniques on any data from the input sources. Hence the user of our system can use advanced modeling techniques to create

⁴ API = Application Programming Interface.

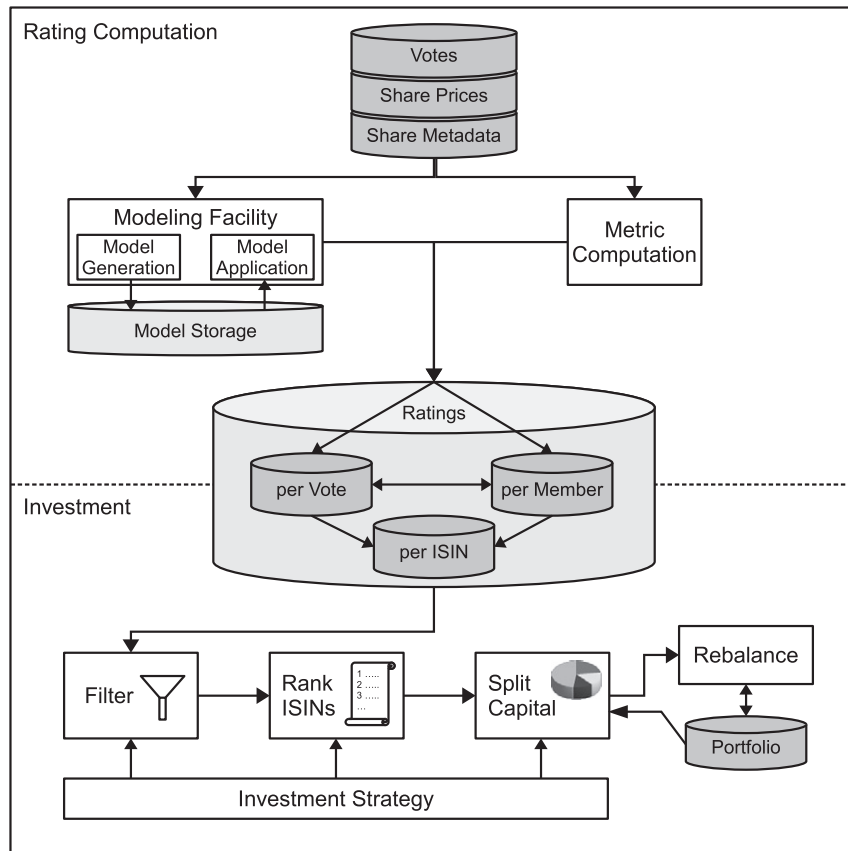


Fig. 2. Overview of the system design.

complex models to support the investment decision [cf. 11,16,20]. When the user wants to create a model, s/he manually conducts a classical train–validate–test cycle on the data accessible by the system to arrive at a model suitable to her/his needs. After creation, a separate model store holds the models created by the user to apply them on new data and use its output as a metric in the rating computation process.

While metrics and models can be computed on any of the levels – vote, crowd member or ISIN level – we ultimately want to arrive at ratings on ISIN level as those are the building blocks for our portfolio. Hence, we need to transform all vote and member metrics into ISIN ratings. To give an example, a member ranking model delivers a metric that indicates the relative quality of a crowd member. Because we need to transform that ranking into an ISIN rating, we weight the votes of each member by the member's rank when aggregating the votes into ISIN ratings such that a vote of a higher ranked member has more weight than a lower ranked one. By doing so, we can combine metrics on different levels into a common metric on ISIN level.

The question of which metrics to combine, and how, depends on the application scenario and is hence subject to the user's experience or trial and error exploration. The system supports this task by making it possible to simulate investment strategies on any metric and thus enables a performance comparison based on a simulated portfolio.

2.3.2. Investment phase

The second phase of the decision support process is the actual investment phase. Despite its purpose to execute a defined investment strategy to manage a portfolio, it also serves to simulate strategies and explore different approaches based on the ratings computed in the previous phase. The simulation approach is similar to that used by Ref. [36].

During the investment phase, the system transforms the stored ratings from the previous phase into a target portfolio. A portfolio in the system consists of a certain cash amount and a number of portfolio positions

representing the shares held in the portfolio. We call the set of parameters defining this transformation an *investment strategy*. It defines which ratings have to be taken into account, how to rank them and how to split the available capital among the ranked ISINs. To modify a portfolio, the system creates and executes buy and sell orders.

During the filter phase, the system selects a subset of the available ISIN ratings as defined by the investment strategy. This filtering involves a diverse set of criteria, e.g. geographical scope, index membership, industry sector or conditions like "at least three crowd votes per rating". Subsequently, the selected set is ordered by a chosen metric which is a component of the rating. If the user requires a ranking by more than one metric, we recommend to add a new metric which combines the other metrics in a defined way so that the decision criteria is transparent and documented for later reference.

After the ranking of ISINs, the available capital needs to be split as defined by the strategy. Example split strategies are: "Split evenly over top 10 ranked ISINs" or "Pick best ranked ISIN of three different industries." The split affects the exposure to any individual ISIN and has to make a good trade-off between potential (focusing on less, but higher ranked ISINs) and risk reduction by diversification (spread capital over many ISINs to reduce correlation between shares). After computing the split, we arrive at a target portfolio composition. The system compares this target portfolio composition to the current portfolio and creates buy or sell orders in order to match the current portfolio to the target composition.

2.3.3. Implementation

We created an artifact to evaluate the functionality of the proposed decision support system design as a prototype in PHP. It has a console interface to enable batch mode operation in a real world scenario, but is also prepared to grant access via a web interface. It connects to a

MySQL database that stores the crowd vote data, the stock quotes, the computed ISIN ratings and the portfolio data.

There are two modules for each of the phases “rating computation” and “investment” which may run either together or independent. The rating computation module expects the time period to compute ratings for as a parameter. The investment module demands a time period, a portfolio id and the ranking metric as parameters. Having two independent modules, the user can experiment with new metrics or modeling approaches while another user can run investment simulations. The modules are coupled via the common rating database. In the current state, the parallel usage of the modules is restricted to different time frames but we can easily solve this issue by duplicating the database, i.e. introducing staging areas.

2.3.4. Modeling facility

For modeling, we provide an interface to the R software package.⁵ We chose this software because it is easy to run in batch mode, has an open architecture including an extensive library for analytical methods and it is available without license cost. For model storage, we use the serializing features of R and store the models in separate files within the host's file system. Upon execution during the rating computation phase, the system takes care to set model parameters as defined by the user, provides the appropriate input data and collects the model results to write it back to the database.

2.3.5. Share price data

A financial website provides the quotes from the Frankfurt Stock Exchange for our system. These data include daily prices (opening, closing, highest, lowest price and transaction volume) and the system uses them as follows: When an order is executed in a simulation run, the price of the share which is bought or sold corresponds to the next trading day's opening quote. This delay provides for the necessary time to collect all VIC votes on a day and then create orders for the next trading day [cf. 22]. A daily system run cycle is not necessarily a restriction to the design, it can also be done weekly or monthly or several times during a day. The length of a reasonable interval depends on the frequency of changes in the crowd vote data set.

3. Prototype

In this section, we show implementation details of our prototype system to provide an application example of the proposed design. We show how to setup a simple investment support task, defining the metric and the investment strategy. Subsequently, we run a simple and a more complex test case on a test period of two years and present the results.

3.1. Definition of performance metric

In this section, we introduce the metric used to identify promising securities (ISINs) based on the wisdom of crowd concept. As mentioned, averaging the judgments of a crowd leads to superior estimates due to bracketing [39]. In terms of investment decisions, averaging several user votes whose target prices over- and underestimate the real target price might lead to a target price which is quite accurate. The metric introduced here uses this effect based on the set of VIC votes for a certain ISIN on a certain day. It is a rather simple metric, so we are able to compute it without the use of the modeling facility.

To quantify the forecast of a vote, we compute the difference in price from the current quote at any given day t to the target price $PRICEGOAL$

at the target date of the vote i for share $ISIN$ with:

$$\Delta p_i(t) = PRICEGOAL_i - PRICE(ISIN, t).$$

If the vote is 100% correct, the price of the share will increase by Δp_i until the target date of the vote.

When it comes to performance, another factor which has to be taken into account is the length of the forecast period. As mentioned, votes can have different runtimes and at any day there will be open votes with a higher or lower amount of remaining days. Logically, we prefer a 2% return in one week over 2% in one month. Earlier returns (which can be reinvested) should be valued higher over later returns of equal amount. We solve that by dividing the estimated price gap of each vote by the remaining runtime in days to arrive at a daily average expected performance:

$$E_i(t) = \frac{\Delta p_i(t)}{\Delta r_i(t)}.$$

Δr_i simply is the remaining runtime in days at time t of vote i which ends at ENDDATE:

$$\Delta r_i(t) = ENDDATE - t.$$

In addition, dividing by the remaining runtime days supports two other purposes: (1) as uncertainty of predictions increases the further they reach into the future, the vote with the shorter and hence more certain time horizons are preferred, given equal absolute prospected returns, and (2) with regard to the specific data set, there are many votes that had no end date specified and hence were closed automatically by the platform after a timeout of 180 days. Hence, every vote that *does* have a specified end date has a runtime that very likely is less than 180 days. So given equal absolute price gaps, votes with specified end date (and hence a shorter runtime than the maximum runtime) are preferred over those open for a longer period. The assumption for doing so is that a more specific vote is likely based on a deeper analysis or superior market insight.

To arrive at a specific investment decision, we roll up those forecasted average daily price changes of individual votes into a joint judgment for an ISIN. For example, if we have two buy and one sell recommendation for the same ISIN, the system combines the individual estimations into a “crowd” judgment by aggregating the individual price estimates, both positive and negative. Now to roll up all the individual estimates per ISIN, we take the mean of the daily expected returns derived for any ISIN from n votes:

$$E_{perf}(t) = \frac{\sum E_i(t)}{n}.$$

Table 1 shows a numerical example of the metric computation on a few sample records as it would turn out on February 5, 2013. The metric is pre-computed and stored in the database for any day and any ISIN appearing in open votes on that day.

Table 1

Calculation example for the rating metric as evaluated on Feb 5, 2013.

Vote	ISIN	End date	Δr (days)	Δp (EUR)	E_i	E_{perf} (ISIN)
1	DE1111111111	10.02.2013	5	4.30	0.86	0.30
2	DE1111111111	05.06.2013	120	−6.00	−0.05	0.30
3	DE1111111111	01.03.2013	24	2.10	0.09	0.30
4	DE2222222222	06.03.2013	29	0.90	0.03	0.04
5	DE2222222222	24.05.2013	108	5.00	0.05	0.04
6	DE3333333333	18.02.2013	13	−3.50	−0.27	−0.07
7	DE3333333333	29.07.2013	174	0.70	0.00	−0.07
8	DE3333333333	30.04.2013	84	4.00	0.05	−0.07

⁵ <http://www.r-project.org/>.

3.2. Filter

We use the filter step to keep only ratings for shares listed in the German Prime Standard Indices, i.e. DAX, MDAX, SDAX, TecDAX – all other shares, including international and penny stocks were excluded from the set. We did that for two reasons: First, this ensures enough liquidity to successfully trade the identified stocks and second, to prevent the possibility of price manipulations by extensive pushing of certain (especially low-priced and low-volume) stocks.

In addition, we filter for ratings with a minimum number of votes (either sell or buy). This is to prevent single extreme votes from distorting the ranking and possibly keep sticking to the top of the list. For the scenario at hand, we require a minimum of three votes for a rating – a deeper investigation on the optimal number of votes might yield further interesting insights.

3.3. Rank, split and rebalance

The system ranks the ISINs by ordering the remaining ISINs by E_{perf} in descending order. This list is already an intermediate result which supports investors in their search of valuable shares using crowd wisdom. Within the system, the next step is to select ISINs from the list and split the available capital to create a target portfolio composition.

For our test runs, we apply a simple and a complex split and rebalancing approach. Please refer to [Section 4.1](#) for details.

4. System evaluation

4.1. Test scenarios

We created two test cases which differ in the way the portfolio is created based on the list of ranked ISINs. Up until the creation of the ranked ISIN list, both test runs are equal.

The first strategy (Test 1) is rather simple, but rational: we invest all cash into the best-ranked ISIN. This is a risky, but also a promising approach. If the forecast is accurate, we maximize our profits by investing in the share with the highest return. We conduct a daily rebalancing of the portfolio: the system checks if the current ISIN in the portfolio is still the number one recommendation of the crowd and if not, it creates and executes a sell order for all of the shares in the portfolio and a buy order for the now best-ranked ISIN. If the portfolio is still empty, the system creates a buy order and invests the total cash position in the best-ranked ISIN. If the best-ranked ISIN is already in the portfolio from previous periods, no trade is initiated.

In a real-world application, investors would strive for a more diversified portfolio to reduce risk by combining shares which are unlikely to correlate in their development. In order to provide such a more sophisticated example of system usage, we demonstrate a second test scenario (Test 2) applying portfolio optimization (PO) on the crowd's Top10 selection of shares. Portfolio optimization, as introduced by Markowitz [30], determines efficient portfolios for a given set of shares. Efficient portfolios either have a maximum return for a given level of risk (measured by standard deviation of historic returns), or a minimum risk for a given level of return. Using quadratic programming, portfolio optimization determines weights for each of the shares in the selection set to arrive at a well-diversified portfolio. More details on portfolio optimization can be found in Refs. [1,30,31]. For our purposes, we use an off-the-shelf implementation of the optimization algorithm in R provided by the *tseries* package [42]. This also demonstrates the system's ability to include more sophisticated modeling facilities by interfacing the R universe with its extensive library of modeling and optimization methods.

To combine the PO with the crowd vote rating, we modified the investment strategy as follows: Instead of taking only the top ranked ISIN from the list, we take the top 10 shares and apply the portfolio optimization on those to determine the optimal weights for a target

portfolio (which can also be 0, i.e. the algorithm may exclude shares). The portfolio optimization algorithm uses historic data of one year to determine the shares' volatility. To create a more common investment scenario, we also increased the rebalancing interval – i.e. the amount of days to pass before the portfolio structure is re-aligned to a given target composition – from one day to 20 (work) days, leading to a monthly re-arrangement of the portfolio structure. Less frequent rebalancing enables the portfolio positions to develop and gain the benefits from a correct investment decision. As less rebalancing means less trades, transaction costs are reduced as well. In addition, we set a minimum position size of 500 EUR (which was not needed in the first test scenario because there was only one position to invest) to avoid creating too small positions which would not be practical due to handling effort and transaction cost efficiency.

This scenario illustrates how the system is able to implement a more complex strategy containing three parameters which provide flexibility to find a suitable investment approach to users: the count of top shares of the ranked list that should be fed into the optimization run, the rebalancing interval in days, and the minimum position size per share.

4.2. Data

The vote data set used for our test run contains 15,727 crowd votes which were created between January 2009 and December 2010 from 1353 distinct members considering 141 different shares. [Table 2](#) summarizes some descriptive statistics.

First, we let the system compute the metric ratings for the whole test period. After that, we conduct a full cycle of the investment phase starting from an empty portfolio with 100,000 Euro cash. The computation of the ISIN ratings for one day takes approximately 3–5 min on the test server (Intel i5-2520M, 2.5 GHz, 4 GB RAM) which is mostly spent on time consuming database queries. The investment simulation phase for Test 1 needs less than a second for a simulated day (i.e. build ranked list, pick first ISIN, build and execute order on simulated portfolio). For Test 2, the R run for the portfolio optimization takes a considerable amount of time (approximately 2–3 min including loading historic data from the database to compute the covariance matrix of share returns), but as we consider longer intervals this also leads to fewer optimization runs and thus to a slightly longer overall runtime than for Test 1.

For transaction cost (TC) calculation, we used the fee structure of a large German broker [14]. The per-trade cost is 4.90 EUR + 0.25% * transaction volume. The maximum cost charged for a transaction is 59.90 EUR (being effective at a transaction volume of and above 22,000 EUR which is the case for every of our simulated transactions). For a more transparent presentation, the transaction costs are not deducted from the cash account of the portfolio

Table 2
Descriptive statistics of vote data set for test runs.

Votes	Absolute	In %
Total	15,727	100%
By market segment:		
• DAX	5647	36%
• MDAX	4589	29%
• SDAX	1815	12%
• TecDAX	3676	23%
By recommendation type:		
• Buy	9169	58%
• Hold	1077	7%
• Sell	2693	17%
• Strong buy	1968	13%
• Strong sell	820	5%
By year:		
• 2009	5976	38%
• 2010	9751	62%
Average votes per day	21.6	
Number of distinct ISINs	141	
Number of members	1353	

(and would hence reduce the subsequent cash amount available for reinvestment), but are rather modeled as being externally funded and thus only affect the final return.

4.3. Results and discussion

In this section, we first describe the results of the rating computation which is equal for both test cases and produces the ranked ISIN list. Then we show and discuss the results of the two test scenarios. We then provide a comparison with two selected public funds and a common risk evaluation of all observed approaches.

4.3.1. Rating computation

For the focal period, an ISIN rating bases on 15.71 votes on average. The top ratings, which ultimately determined the shares selected for the portfolio in Test 1, are based on 10.30 votes on average. This shows that not necessarily the number of votes determines the winning position. The average number of open votes for a day was 123.24 (Fig. 3 shows the development over time) which means, on average, 123 ISINs were evaluated for a day. In a by-year comparison, this figure is rather low in 2009 with 110.53 and goes up to 136.05 in 2010. This shows that the analyst coverage increased due to increasing activity on the platform indicated by the growing number of user votes (cf. Table 2).

4.3.2. Results of Test 1

Fig. 4 illustrates the investment simulation results for Test 1. The graph shows a standardized performance comparison of the portfolio performance, the portfolio performance after deduction of transaction costs and the performance of the DAX German market index (buy-and-hold). The lines indicate how an initial investment of 100,000 EUR develops over time according to the respective performance. Additionally, the chart shows the accumulated transaction cost which decrease the profits from the investment as can be seen by the spread between the two portfolio lines.

Table 3 provides an overall performance comparison which shows that the system outperforms a DAX index buy-and-hold strategy to a large extent. Even after considering the transaction cost, the performance is still roughly twice as large as the benchmark. In a daily comparison, on 275 of 520 days (53%) the crowd vote outperforms the DAX index benchmark. Interestingly, only 75 ISINs of the total of 141 (53%) appeared in the first ranked positions at least once; the rest never made it to the top position.

However, a closer look at Fig. 4 reveals that the performance comparison can roughly be split in two phases: While the performance of

the crowd votes is clearly underperforming the DAX index in 2009, the relation reverses in 2010 when the portfolio starts to outperform the benchmark index. A possible explanation of why the performance increases in 2010 might be the increasing activity on the platform leading to a higher stability in crowd votes. As stated, Fig. 3 indicates that the number of open votes for each day stabilizes on a high level in 2010. At the same time, the number of monthly trades triggered by the system enters a downward trend, also indicating a more stable crowd judgment leading to less fluctuation in the portfolio, even when checking daily for the single highest performing stock. These observations are only indicative and would need further analyses which are beyond the scope of this paper.

4.3.3. Results of Test 2

In Test 2 we apply a more sophisticated portfolio optimization (PO) on the crowd's Top10 selection of shares. This is a more realistic setup and depicts a scenario that might also be found in business practice.

Fig. 5 shows that the diversified PO portfolio still outperforms the market index and reaches a performance of 70 points (without TC) or 59.2 points (including TC), respectively, above the benchmark (see Table 4). In comparison to Test 1, the overall performance of the portfolio is slightly lower, yielding only a return of 110% compared to 123% in Test 1. Interestingly, after deduction of transaction cost, Test 2 portfolio still shows a return of 99.8% while the performance of the Test 1 portfolio drops down to 88.5%. This is caused by the fact that in Test 2 the portfolio is rebalanced on a monthly basis compared to a daily rebalancing in Test 1 and thus the transaction costs are lower in Test 2.

The Test 2 portfolio outperformed the index benchmark on 265 of 522 days or 51%, respectively. The mean amount of positions in the portfolio is 4.77. This shows that the portfolio optimization algorithm is quite selective and reduces the number of possible positions on average by more than half (from 10 possible values). This includes the effect of a minimum position size of 500 EUR which prevents very low weights from the optimization run to result in a portfolio position.

4.4. Performance comparison with public funds

We compare the system results with the performance of two selected public funds over the same period. Such funds provide another benchmark for the performance of the two test runs and give a reference to evaluate the external validity of the system results. As public funds are managed by professional institutions with a defined investment strategy and a specific analytic approach, they can serve as a

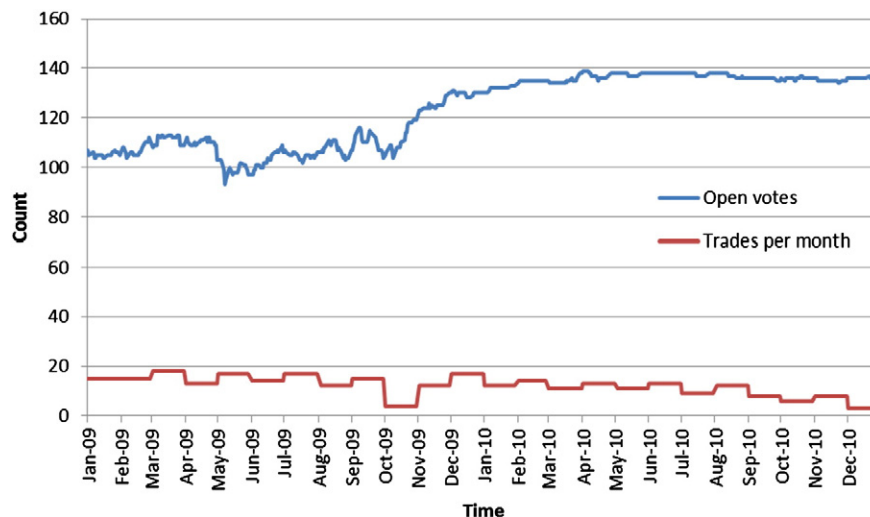


Fig. 3. Development of the number of open votes and monthly trade volume.

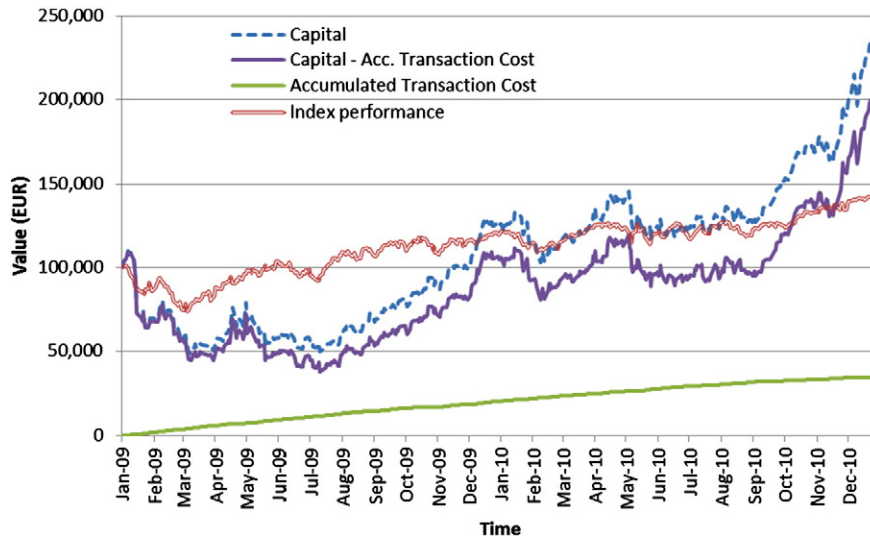


Fig. 4. Results of Test 1 from January 2009 to December 2010.

state-of-the-art benchmark for our DSS which tries to solve the same task.

To have a comparable benchmark, we have to choose funds that meet some requirements: the portfolio must pick shares from a similar investment universe (i.e. German Prime Standard), provide sufficient historic data for our test period and should be representative for their kind, i.e. have a high reputation or be important in terms of size. Based on these prerequisites, we chose the following two funds for comparison:

- Fund 1: DWS Deutschland (ISIN DE0008490962): One of the largest funds for German Standard shares (3.3 billion EUR in assets), Morningstar Silver Rating [35]. The strategy of portfolio building is a mixture of top-down elements (e.g. sector allocation, share of mid/small caps) and bottom-up (strong emphasis on fundamental analysis using standard valuation models, management briefings and common stock metrics (P/E & P/B ratio, EBITDA), among others). The investment ratio depends on technical indicators, sentiment and general assumptions about the market development [35]. This fund is representative for a large, traditionally managed standard fund.
- Fund 2: Allianz Thesaurus AT EUR (ISIN DE0008475013): This fund (153 million EUR in assets) is based on a momentum model that tries to identify trends to over- or underweight certain shares from the investment universe. The model also considers risk contribution of shares [13]. We chose this fund to compare our test results to a professional approach based on quantitative modeling.

Fig. 6 shows that the two public funds have approximately the same performance at the beginning of our observation period, but start to diverge at the end of 2009, with Fund 1 performing superior until the end of the test period. The Test 1 portfolio clearly underperforms all alternative portfolios in 2009, but rises up to the top until the end of 2010. The optimized portfolio of Test 2 shows a better and steadier performance, climbing to the top early and staying there almost until the end of the period. However, the performance differences might be

caused by different portfolio risks which we will hence examine in the following section.

4.5. Risk evaluation

A common measure for the risk associated with the investment into a certain security is the volatility of its returns. Hence, to judge the risk of a security, we compute the standard deviation of security prices [cf. 24,38]. When two securities yield the same return, the one with the lower volatility had a more steady development and as such a lower risk. Usually, higher return potential comes with a higher risk. To measure the “price of risk” and make investments which differ both in potential and risk comparable to one another, Sharpe [37] introduced the “Reward-to-Variability-Ratio” (Sharpe Ratio). This ratio S of an investment is defined as:

$$S = \frac{\bar{D}}{\sigma_D}.$$

\bar{D} is the average outperformance against a riskless investment over a time period divided by the standard deviation σ_D of these outperformance returns during the time period [38]. The Sharpe Ratio indicates the reward for each unit of risk and makes different investments comparable. A higher Sharpe Ratio indicates a better return-to-risk relation and hence a better investment. In our work, we assume there is no risk-free alternative to stock investment (we force all cash to be invested into shares) in order to focus on examining the quality of the stock selection process, i.e. we use absolute returns for the Sharpe Ratio.

Table 5 summarizes the Sharpe Ratio and its components for all portfolios examined in this paper. We consider all portfolio performances before transaction costs because we do not have information on the transaction costs of the public funds and for replicating the DAX benchmark. Comparing the results of Test 1 and Test 2, we see that the mean daily return of Test 1 is higher, but this also holds for the standard deviation and hence — volatility or risk. The Sharpe Ratio of Test 1 portfolio (0.0648) is lower than those of Test 2 (0.0791) which means that the PO was a valuable strategy, because it reduced risk more than it reduced returns. Looking at Fig. 6, we see that the Test 2 portfolio avoids the low trough of the Test 1 portfolio in 2009, but also exhibits more side-ward development later on while Test 1 portfolio increases rapidly in 2010. Fig. 6 thus nicely illustrates the reduced volatility of the Test 2 portfolio.

The public fund 1 has a lower return than the system portfolios, but also a lower risk, leading to a Sharpe Ratio of 0.0604 which is above the

Table 3

Test 1 results: comparison of benchmark and portfolio performance incl. transaction costs.

Instrument	Start January 2009 (EUR)	End December 2010 (EUR)	Absolute return (EUR)	Return in %
DAX benchmark	100,000	140,552	40,552	40.6%
Portfolio	100,000	223,164	123,164	123.2%
Portfolio-TC	100,000	188,484	88,484	88.5%

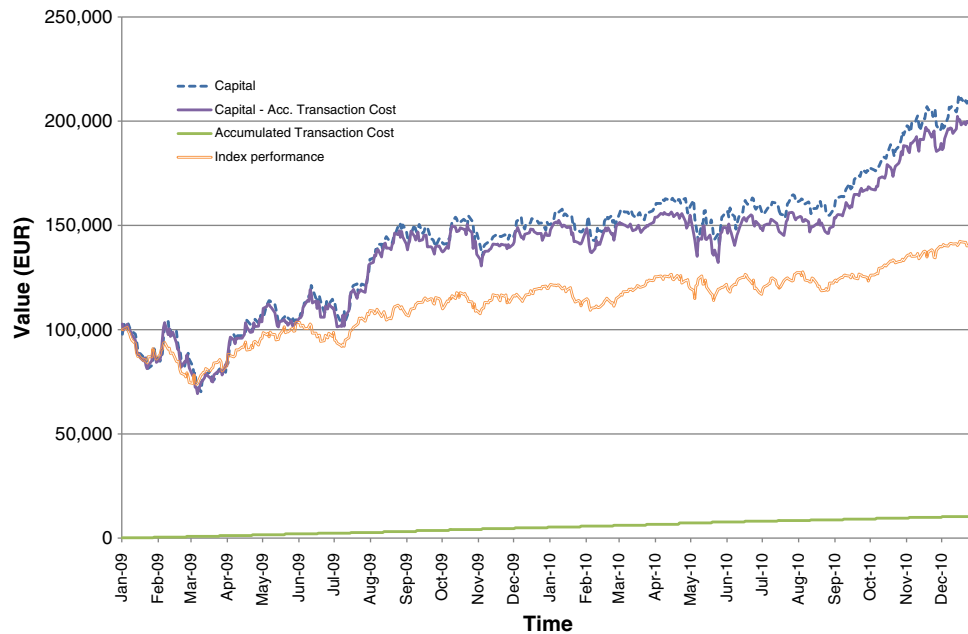


Fig. 5. Results of Test 2 from January 2009 to December 2010.

benchmark, but still below the two system portfolios. Hence, the return loss is not completely compensated by additional reliability in comparison to the test portfolios of the system.

Public fund 2 shows a worse performance for the test period for both return and risk, even below the benchmark. It shows for the test period that an approach based on a quantitative model implementing a momentum strategy performs below the results of our test runs.

Both public funds have a lower risk for the test period, indicating the strict risk management of public funds compared to our simple test approaches. In addition, both funds tend to a high share of large caps [13,35] which make them stick close to the DAX benchmark (which contains the 30 largest public companies of Germany). In contrast, our system's approach did not prefer large caps over small or mid caps but let alone the crowd's estimate determine preferences. If we focus however on absolute returns and/or on the Reward-to-Variability-Ratio, our Test scenarios, especially Test scenario 2, clearly outperform all other benchmarks.

5. Conclusion

Research shows many approaches for investment decision support systems but none of them enables users to take advantage of the wisdom of crowd even though it has been shown to be advantageous to investment decisions [4,22] for the price of complex analytical effort. We address this gap and propose the presented system design; it provides flexible means to transform raw crowd vote data into actionable investment decisions, up to automatic portfolio maintenance. By building a prototype and running two test cases that clearly outperformed a market benchmark index over a period of two years, we showed the viability of our concept. Even with a rather simple strategy, we achieve a portfolio performance of 123%, outperforming the market benchmark by 83 points

and still 48 points after transaction costs. A second test case illustrated how a sophisticated optimization approach can be included to consider risk and build a diversified portfolio that still outperforms the market benchmark by 70 or 60 points (after TC), respectively. A comparison with public funds shows that the results are reasonable given the institutional boundaries of such funds compared to the prototypical character of our test scenarios. Further steps have to be conducted before a real-world application of the system is deemed appropriate. Specifically, further testing in different market phases and an appropriate strategy formulation as well as adaption to real-world restrictions in portfolio building is necessary. Minimum position size and adjustable rebalance intervals are already first steps in fulfilling those requirements.

Both cases show that the proposed decision support system enables investors to make use of the wisdom of crowd and automatically test and apply strategies based on this source. The results of our system can either be used standalone or be integrated with other traditional measures from fundamental or chart analysis. It could serve as an additional signal when it comes to an investment decision with respect to a certain stock.

From a scientific perspective, the system provides a laboratory environment to further analyze the phenomenon of crowd wisdom in financial markets. The infrastructure provided by the system enables us to further investigate and simulate promising models and investment strategies. One possibility to arrive at the best strategy would be to create numerous software agents that try to maximize their outcome by using the proposed DSS (cf. [23]). Such a multi-agent simulation would help to increase our understanding on the exploitation of the wisdom of the crowd in financial markets.

Regarding future improvements of the proposed system could include the introduction of target risk to guide the decisions proposed by the system and match them with an investor's risk profile. Additionally, literature suggests that the predictive power of crowds can be increased by selecting a group of experts based on historical performance within the crowd [22]. This could be taken into account by creating a ranking metric that gives more weight to potentially more experienced crowd members. The modeling facility of our proposed system already allows incorporating the member experience in the ranking models. However, since this approach contradicts the wisdom of crowd concept which strives for maximum diversity [41], it is a different approach than used in the application example of this paper, but is supported by the flexibility of the system design and worth further exploration.

Table 4
Test 2 results: comparison of benchmark and portfolio performance incl. transaction costs.

Instrument	Start January 2009 (EUR)	End December 2010 (EUR)	Absolute return (EUR)	Return in %
DAX benchmark	100,000	140,552	40,552	40.6%
Portfolio	100,000	210,600	110,600	110.6%
Portfolio-TC	100,000	199,824	99,824	99.8%

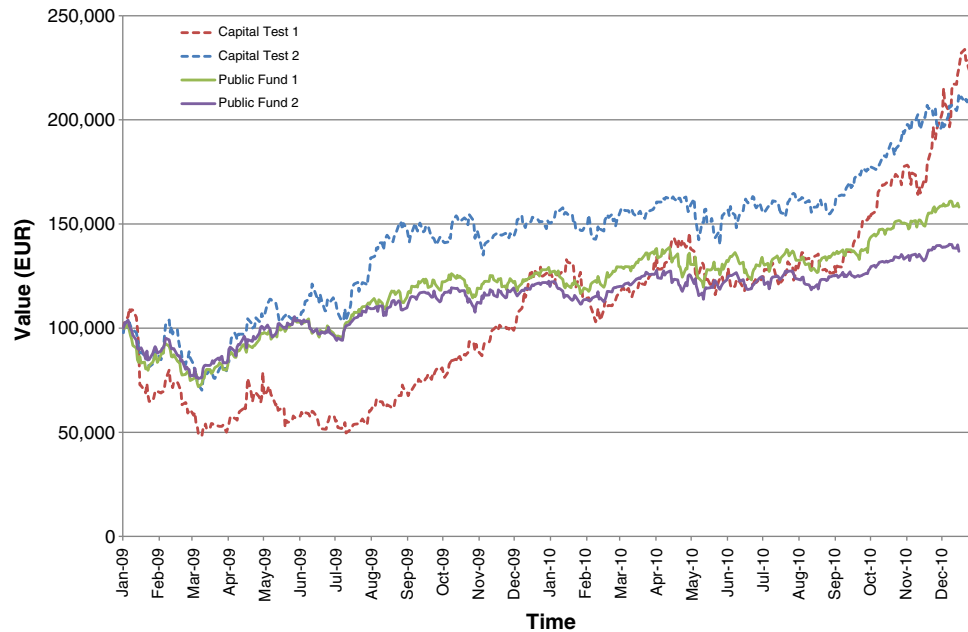


Fig. 6. Comparison of test portfolios and public funds.

Table 5

Mean and standard deviation of daily returns, Sharpe Ratio for all observed portfolios (Jan 2009–Dec 2010).

Portfolio	Mean of daily return	Standard deviation of daily return	Sharpe Ratio
DAX benchmark	0.0009	0.0150	0.0581
Portfolio Test 1	0.0021	0.0321	0.0648
Portfolio Test 2	0.0016	0.0208	0.0791
Public fund 1	0.0011	0.0174	0.0604
Public fund 2	0.0007	0.0155	0.0476

Acknowledgments

The authors thank Theresa Krimm, Michael Nofer and Renata De Sousa for their comments and support on improving the paper.

References

- [1] M. Andrecut, Portfolio Optimization in R, arXiv, 12013.
- [2] W. Antweiler, M.Z. Frank, Is all that talk just noise? The information content of internet stock message boards, *The Journal of Finance* 59 (2004) 1259–1294.
- [3] K.J. Arrow, R. Forsythe, M. Gorham, R. Hahn, R. Hanson, J.O. Ledyard, et al., The promise of prediction markets, *Science* (New York, N.Y.) 320 (2008) 877–878.
- [4] C. Avery, J. Chevalier, R. Zechhauser, The “CAPS” Prediction System and Stock Market Returns, 2011.
- [5] N. Baba, N. Inoue, Utilization of soft computing techniques for constructing reliable decision support systems for dealing stocks, *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02* (Cat. No.02CH37290), IEEE, 2002, pp. 2150–2155.
- [6] P. Beraldi, A. Violi, F. De Simone, A decision support system for strategic asset allocation, *Decision Support Systems* 51 (2011) 549–561.
- [7] J.E. Berg, T.A. Rietz, Prediction markets as decision support systems, *Information Systems Frontiers* 5 (2003) 79–93.
- [8] J. Bollen, H. Mao, X. Zeng, Twitter mood predicts the stock market, *Journal of Computational Science* 2 (2011) 1–8.
- [9] M.M. Carhart, On persistence in mutual fund performance, *The Journal of Finance* 52 (1997) 57–82.
- [10] L. Chen, P. Goes, J.R. Marsden, Z. Zhang, Design and use of preference markets for evaluation of early stage technologies, *Journal of Management Information Systems* 26 (2009) 45–70.
- [11] V. Cho, MISMS – a comprehensive decision support system for stock market investment, *Knowledge-Based Systems* 23 (2010) 626–633.
- [12] S. Chou, C. Yang, C. Chen, F. Lai, A rule-based neural stock trading decision support system, *Computational Intelligence for Financial Engineering, 1996. Proceedings of the IEEE/IAFE 1996 Conference On, 1996*, pp. 148–154.
- [13] B. Claus, Morningstar Research Report Allianz Thesaurus AT EUR, 2013. (September 20, 2012).
- [14] comdirect Bank AG, Preis- und Leistungsverzeichnis, 2013. 5.
- [15] E. Dahan, A. Soukhroukova, M. Spann, New product development 2.0: preference markets—how scalable securities markets identify winning product concepts and attributes, *Journal of Product Innovation Management* 27 (2010) 937–954.
- [16] J. Dong, H.S. Du, S. Wang, K. Chen, X. Deng, A framework of Web-based Decision Support Systems for portfolio selection with OLAP and PVM, *Decision Support Systems* 37 (2004) 367–376.
- [17] E.F. Fama, K.R. French, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33 (1993) 3–56.
- [18] R. Forsythe, F. Nelson, G. Neumann, J. Wright, Anatomy of an experimental political stock market, *The American Economic Review* 82 (1992) 1142–1161.
- [19] J. Giles, Internet encyclopaedias go head to head, *Nature* 438 (2005) 900–901.
- [20] S.S. Groth, J. Muntermann, Supporting investment management processes with machine learning techniques, *Wirtschaftsinformatik Proceedings, Paper, 107*, 2009.
- [21] A.R. Hevner, S.T. March, J. Park, S. Ram, Design science in information systems research, *Mis Quarterly* 28 (2004) 75–105.
- [22] S. Hill, N. Ready-Campbell, Expert stock picker: the wisdom of (experts in) crowds, *International Journal of Electronic Commerce* 15 (2011) 73–102.
- [23] O. Hinz, M. Spann, Managing information diffusion in Name-Your-Own-Price auctions, *Decision Support Systems* 49 (2010) 474–485.
- [24] J.C. Hull, *Risk Management and Financial Institutions*, Third ed. Pearson Prentice Hall, New Jersey, 2007.
- [25] R. Kuo, C. Chen, Y. Hwang, An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network, *Fuzzy Sets and Systems* 118 (2001) 21–45.
- [26] R.P. Larrick, J.B. Soll, Intuitions about combining opinions: misappreciation of the averaging principle, *Management Science* 52 (2006) 111–127.
- [27] N. Liu, K. Lee, An intelligent business advisor system for stock investment, *Expert Systems* 14 (1997).
- [28] I. Lorge, D. Fox, J. Davitz, M. Brenner, A survey of studies contrasting the quality of group performance and individual performance, 1920–1957, *Psychological Bulletin* 59 (1958) 257–272.
- [29] B.M. Lucey, M. Dowling, The role of feelings in investor decision-making, *Journal of Economic Surveys* 19 (2005) 211–237.
- [30] H. Markowitz, Portfolio selection, *The Journal of Finance* 7 (1952) 77–91.
- [31] H.M. Markowitz, Foundations of Portfolio Theory, *Journal of Finance* XLVI (1991) 469–477.
- [32] G.P. Moynihan, P. Purushothaman, R.W. McLeod, W.G. Nichols, DSSALM: a decision support system for asset and liability management, *Decision Support Systems* 33 (2002) 23–38.
- [33] J. Muntermann, Towards ubiquitous information supply for individual investors: a decision support system design, *Decision Support Systems* 47 (2009) 82–92.
- [34] M. Nofer, O. Hinz, Are Crowds on the Internet Wiser than Experts? – The Case of a Stock Prediction Community, 2013.
- [35] S. Nöth, Morningstar Research Report DWS Deutschland, 2013. (July 19, 2013).
- [36] E.J. Ruiz, V. Hristidis, C. Castillo, A. Gionis, A. Jaimes, Correlating financial time series with micro-blogging activity, *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining – WSDM'12, 2012*, p. 513.
- [37] W.F. Sharpe, Mutual fund performance, *The Journal of Business* 39 (1966) 119–138.
- [38] W.F. Sharpe, Asset allocation: management style and performance measurement, *The Journal of Portfolio Management* 18 (1992) 7–19.

- [39] J.B. Soll, R.P. Larrick, Strategies for revising judgment: how (and how well) people use others' opinions, *Journal of Experimental Psychology. Learning, Memory, and Cognition* 35 (2009) 780–805.
- [40] M. Spann, B. Skiera, Internet-based virtual stock markets for business forecasting, *Management Science* 49 (2003) 1310–1326.
- [41] J. Surowiecki, *The Wisdom of Crowds*, Anchor Books Random House, New York, 2005.
- [42] A. Trapletti, K. Hornik, *tsseries: Time Series Analysis and Computational Finance*, 2013.
- [43] R. Tsaih, Y. Hsu, C.C. Lai, Forecasting S&P 500 stock index futures with a hybrid AI system, *Decision Support Systems* 23 (1998) 161–174.



Jörg Gottschlich studied business administration and computer science at the University of Bamberg, focusing on information systems, database applications and management and received his diploma degree (equiv. master degree) in 2008. He wrote his diploma thesis about "Continuous improvement of cross-selling offers by automatic analysis of customer reactions" and received the "Excellent Research" audience award at SAS Forum conference in 2008. Since then, Jörg worked for the international strategy consultancy A.T. Kearney where he provided fact-based decision support for international strategic projects in different industries (e.g. Finance, Retail, and Manufacturing). Jörg joined the Chair of Information Systems at the TU Darmstadt in 2011 as Research Assistant and Ph.D. student. His work

includes research on decision support and recommender systems and social networks.



Oliver Hinz studied at the TU Darmstadt Business Administration and Information Systems with main focus on Marketing, Software Engineering and Computer Graphics. After receiving his diploma (equiv. master degree) he worked several years for the Dresdner Bank as a consultant for business logic. Oliver started working as a Research Assistant in March 2004 at the Chair of Electronic Commerce and received his Ph.D. in October 2007. Oliver Hinz joined the Marshall School of Business (University of Southern California) as visiting scholar for 4 months and received the SinnerSchrader stipend for young researchers for 2007. He has also been awarded with the dissertation prize of the Alcatel-Lucent-Stiftung 2008, the Erich-Gutenberg-Prize 2008 and the science prize

"Retailing 2009" of the EHI Retail Institute. He is also the winner of the honorable Schmalenbach prize for young researchers in 2008. He supported the E-Finance Lab as Assistant Professor for E-Finance & Electronic Markets, joined the TU Darmstadt in April 2011 and heads the Chair of Information Systems | Electronic Markets.

Oliver's research has been published or is forthcoming in journals like *Information System Research (ISR)*, *Management Information Systems Quarterly (MISQ)*, *Journal of Marketing (JM)*, *Journal of Management Information Systems (JMIS)*, *International Journal of Electronic Commerce (IJEC)*, *European Journal of Operational Research (EJOR)*, *Decision Support Systems (DSS)*, *Electronic Markets (EM)*, *Business & Information Systems Engineering (BISE)* and in a number of proceedings (e.g. *ICIS*, *ECIS*, *PACIS*).