Hermes: Guidance-enriched Visual Analytics for economic network exploration

Roger A. Leite [a,*], Alessio Arleo [a], Johannes Sorger [b], Theresia Gschwandtner [a], Silvia Miksch [a]

[a] *Vienna University of Technology - TU Wien, Austria*
[b] *Complexity Science Hub Vienna, Austria*

ARTICLE INFO

ABSTRACT

The economy of a country can be modeled as a complex system in which several players buy and sell goods from each other. By analyzing the investment flows, it is possible to reconstruct the supply chain for the production of most goods, whose understanding is important to analysts and public officials interested in creating and evaluating strategies for informed and strategic decision making, for instance, adjusting tax policies. Those networks of players and investments, however, tend to be complex and very dense, which leads to over-plotted visualizations that obfuscate precious information such as the dependencies between productive sectors and regions. In this paper, we propose *Hermes*, a guidance-enriched Visual Analytics environment (named after the Greek God of Commerce) for the exploration of complex economic networks, to uncover supply chains, regions' productivity, and sector-to-sector relationships. With practical knowledge regarding guidance, we designed and implemented a visual sub-graph querying approach to extract patterns from such complex investment graphs obtained from real-world data.

We present a three-fold evaluation of the system: we perform a qualitative evaluation of our approach with three domain experts, a separate assessment of the proposed guidance features with an expert researcher in this field, and a case study of *Hermes* using a bank account network dataset to demonstrate the generalizability of our approach.

1. Introduction

Understanding the behavior of the players in an economic network is a challenging and profitable problem. The obtained knowledge could be used to maintain a sustainable environment to support the country's growth. For example, understanding how in a country productive sectors and regions are related allows the government to plan strategic policies that support positive economic trends. Typical tasks involve looking for weak/strong connections, supply dependencies, or growing and shrinking investment trends. While Artificial Intelligence (AI) is very useful to detect known patterns as well as expected and non-expected outliers, when it comes to reasoning and exploration of a data set, a Visual Analytics (VA) tool works as a brush to an artist. By coupling computational power with the power of the human visual system, we enable domain experts to derive insights which would not be possible with other approaches.

In this work, we focus on the problem of exploring country-wide complex financial networks, aiming at unveiling investment networks, supply chains, and economical links between productive sectors and geographical regions. We present *Hermes*, a VA exploration environment for the exploration of economic networks featuring **time-oriented** and **multivariate** data. Our system is composed of an "Overview" board to provide a high level analysis of the available data and a "Network" board for details on demand. In the latter we propose a visual approach for the definition of patterns to be searched within the network that can be defined by creating and combining "profiles", that are groups of nodes that comply to a set of user-defined constraints. We also include guidance-enriched (Ceneda et al., 2017, 2019) components to aid the users in the definition of queries and profiles, to enable a fully interactive visual approach to profile creation, pattern drawing, and pattern matching. Therefore, our main contributions are:

* Corresponding author.
E-mail addresses: roger.leite@tuwien.ac.at (R. A. Leite), alessio.arleo@tuwien.ac.at (A. Arleo), sorger@csh.ac.at (J. Sorger), theresia.gschwandtner@tuwien.ac.at (T. Gschwandtner), silvia.miksch@tuwien.ac.at (S. Miksch).

- We designed a **visual support for the definition of designated profiles**, which can summarize multiple node types and facilitate a compact construction of query subgraphs.
- We integrated a **guidance-enriched visual subgraph querying approach** that aids the analyst in building subgraph queries from the overabundance of possible node and edge combinations.
- We conducted a **three-fold evaluation**: (1) a task-based qualitative user study with three domain experts collecting insights and guidance update rate measurements, (2) the assessment of our guidance design by an expert in guidance methods for VA, and (3) a use case description demonstrating the generalizability and applicability of *Hermes* to a second domain.

2. Related work

Hermes can be framed as a VA approach for the analysis of economic networks using a visual language to perform graph queries with guidance elements. Therefore, we describe the related work by grouping in four categories.

VA for Financial Domain: Visualization proved to be prolific in the financial domain. FinanceVis (Dumas et al., 2014) presents a survey involving over 50 research articles, totaling over 70 different techniques, related to the presentation and visualization of financial data. It also features an online collection of the surveyed papers for easy access. The paper concludes that visualization could empower analysts in the financial domain, and encourages further collaboration. Ko et al. (2016) focused on gathering expert feedback and opinions on existing VA approaches in finance. Interviews were also targeted at understanding possible improvements and preferences about visual abstractions, interaction methods, automated techniques, and data sources. When designing *Hermes*, we drew inspiration from the several visualization approaches in the financial domain Dumas et al. (2014), and from the strong motivation and appeal for more real-world based solutions (Ko et al., 2016).

Network Analysis in Finance: To enable the VA process on financial networks, graph drawing and visualization played a major role, with node-link representations being often used as an intuitive way to present such networks to the users. Allen and Babus (2009) argue that node-link graph diagrams are a natural solution for the representation of financial systems and efficiently explain certain economic phenomena such as resources exchange. This directly supports part of our approach design choices. Tekušová and Kohlhammer (2008) visually supports the users in understanding complex shareholding networks. It incorporates different economic metrics and supports the identification of ultimate shareholders. The identification of network players is a common ground of this work and *Hermes* (see Section 3.3): R4. VisFAN (Didimo et al., 2011) deals with the observation financial transactions between entities to identify money laundering patterns and potential frauds. The system also allows to cluster and aggregate elements for bottom-up exploration with both automatic and manual clustering techniques. Different from this paper, *Hermes* can group elements by the profile creation feature (see Section 4.3.2: A). Huang et al. (2009) present an event monitoring approach for stock markets security. It features a linked-views approach with node-link diagrams for network analysis and a 3D treemap for market performance visualization. The main goal of this approach is to identify the attackers (the sources of the fraud), and further attack plans, while reducing the amount of false-positives generated by AI techniques.

Origraph (Bigelow et al., 2019) supports network data creation, reshaping, and filtering. The tool allows the user to execute these operations with little to no programming according to the level of details required by the task. In *Hermes*, we were inspired by this work to provide for users with a non-programming approach to generate queries. Differently from Origraph, however, *Hermes* allows to aggregate filters together, to generate profiles that fulfill different constraints. Fujiwara et al. (2018) presented a solution to automatically compose a network visual summary based on a rank model (V2D2) and a reduction algorithm. The rank model orders the importance of interactions and provides dependencies among them. Even lacking user studies, the proposal automatic approach proved to be efficient for preparing large networks for analysis. Focused on non-expert users, Van den Elzen and Van Wijk (2014) navigates from network detail to overview by network elements selection and aggregation. In *Hermes*, we follow a similar approach, providing users with different boards that provide different perspectives of the data. As a limitation, they claim to present visual clutter issues with large selections (above 10) due to its clustering or community detection algorithms.

Graphical Queries on Graphs: A typical task in the analysis of financial network is to search for pre-defined patterns and subgraphs, to highlight some transfers and behaviors of interest (e.g., money laundering, tax evasion, supply chain management). To allow users to easily define these patterns, research on VA for finance branched into systems that incorporated visual languages allowing the definition of patterns and profiles of interest, which are used to query the original, complex network. An example of this approach is provided by Didimo et al. (2018). By using an already established visual language for pattern definition (Didimo et al., 2015), they design a system for highlighting malicious transactions between taxpayers (that includes both individuals and legal entities). Differently from our approach, the system is designed to operate in an application domain, while we propose *Hermes* as a more flexible approach more oriented on investments rather than individual transactions. Koenig et al. (2010) discuss about the visualization of graph query search results in attributed graphs. This means that each node belongs to a specific category derived by its attributes (e.g., movies, actors, directors). Similarly to our approach, the users define the patterns by combining the different node categories (represented as nodes) together. To display the results, the obtained nodes are mapped to the user-defined patterns, which helps users to keep their mental maps. On the other hand, this approach limits the number of data items that can be shown in the node-link diagram of query results. The problem of designing a suitable visual query language for graph datasets was also discussed by Chau et al. in their GRAPHITE (Chau et al., 2008) system and, in more depth, by Pienta et al. with their VISAGE (Pienta et al., 2016). Other than providing a platform and visualization for querying graphs, the paper introduces a graph "auto-complete" feature to aid the user to construct and refine queries. This inspired us to include in *Hermes* specific features to assist the users in building the queries. To do so, we focused on the use of "Guidance", a relatively modern field of research that has been recently characterized for VA (Ceneda et al., 2017).

OntoVis (Shen et al., 2006) allows the analysis of social networks by semantic and structural abstraction. It aims at understanding who are the main actors and what are the common behaviors of a social network. Structural abstraction is achieved by the topological information such as node degree and connectivity while semantic abstraction is achieved by filtering nodes that belong to the same type. Different from the "profiles" presented by *Hermes*, that allows for fine-tuned group filters involving different numerical and categorical attributes, the semantic abstraction presented by OntoVis is limited by the existing node types.

Guidance: Guidance has roots in the field of Human Computer Interaction (HCI) (Horvitz, 1999), but Ceneda et al. (2017) extended van Wijk's model of visualization to integrate guidance

into the VA process. They use existing guidance approaches from the literature to illustrate the proposed model. Ceneda et al. (2018a,b) designed and implemented a data-driven guidance tool to demonstrate that guidance techniques have the potential to improve data analysis. Most recently, Ceneda et al. (2019) conducted an elaborated literature research to provide an overview of how guidance is tackled in different approaches. In this work, the authors highlighted the importance of knowing the actual state of the analysis. We designed *Hermes*' guidance features with the same "context-awareness" principle.

3. *Hermes* design context

The main goal of this section is to present the design context of *Hermes*, as well as to describe design and implementation choices.

3.1. Data, user, and tasks

We designed and developed our approach based on the design triangle methodology (Miksch and Aigner, 2014) and Munzner's nested model (Munzner, 2009), by characterizing the problem at hand according to the underlying data, the involved users, and the tasks they have to carry out.

Data. The economic data that our target users analyze can be separated into two categories: the first one describes aggregate information, e.g., on national level, that report the monetary in & out flows between industry sectors within a fiscal year ("IO Tables" Timmer et al., 2015). A transaction in such a table describes, for example, the amount of investments (in millions of Euros) that went from the "Agriculture" sector to the "Manufacturing" sector. Sectors are encoded according to the ÖNACE specification (ÖNACE, 2019). The second data category complements the aggregate information of IO Tables with more granular/regional details, e.g., on the level of individual firms within a region of a country. For this category we used the "Sabina" dataset (Wien, 2019): it contains the balance sheets of all the firms in Austria, including cash flows, locations, branches, and sectors. The combination of both datasets (see Section 4.1) creates the **Exchange Supply Chain (ESC)** dataset which was used during the design and implementation of *Hermes*. The dataset features 172 nodes and 20,709 edges.

Users. *Hermes* target users are financial and economical experts involved in the development of regional policies and investment plans, for instance in order to economically revitalize a region. Users are expected to be proficient in statistical evaluation of tabular data, but are not expected to have any experience in visual analysis or network analysis.

Tasks. The task that *Hermes* aims at supporting, is to improve the understanding of national economic investment networks based on the integration of aggregated and granular data. After discussing with our domain experts collaborators (not involved on the evaluation), we could identified four sub-tasks: the exploration of the economic flow network existing between sectors (**T1**) and regions (**T2**), the evaluation of how sectors relate and impact each region (**T3**), and supply chain exploration (**T4**) (e.g., predominantly buying/selling regions). The identification of these four sub-tasks support us to better understand the potential of our collaboration with the domain experts.

The combination of the acquired knowledge about data, user, and tasks support us to define the Research Challenges 3.2.

3.2. Research challenges

In the context of *Hermes*, we are dealing with several challenges that result from the underlying data as well as the tasks that should be enabled on them. Understanding the research challenges is an important step on the solution design. It is a broader view of the problem that helps us to plan, measure efforts, derive more specific requirements (see Section 3.3), and derive future work suggestion.

Dense Dynamic Networks: Large economic networks are dense and dynamic, with large and frequent fluctuations. While automatic approaches can still be used to detect outliers and confirm known patterns, VA techniques support economic networks analysis, as they enable users to interactively drill down on the dense data with multiple views, improving the exploration process.

Multi-scale Context: the analysis and exploration of transactions on and across different scales, i.e., regional, national, global, requires visual and analytical scalable solutions.

Geo-political Context: To better understand transactions, we need to consider the motivations that guide them. It is known that geopolitical, geographical, and economical contexts influence this behavior, e.g., buying from a local or a distant supplier, or the presence of a quarry or oil field. Considering the ever changing local and global scenarios, an analysis tool must be flexible in conveying these different contexts.

Temporal Trends: Economic network analysis involves the observation of stable behavior over time (e.g., distinguishing growing regional sectors from shrinking and stable ones). However, there are many aspects (scale, scope, arrangement, viewpoint) concerning time-oriented analysis that need to be analyzed efficiently due to their complex data features (Aigner et al., 2011).

3.3. Requirements

By combining the tasks derived by the problem understanding (see Section 3.1), the broader research challenges context (see Section 3.2), and with further collaboration of three economists, we derived four main requirements for our system. The economists that collaborated during the requirements and the design development were not participants of the evaluation presented in Section 5. By defining a set of requirements, we can also give a better justification for our design choices and provide a robust framework to design our evaluation.

R1: Overview of region and sector distributions. An economic environment is in constant evolution: the transactions between its major players (the companies) can leave a significant and lasting impact on the economy of the region they belong to as well as their foundation and demise. An overview of these aspects is important to understand the main sectors of an economy and supports insights such as the most and least prominent sectors among different regions.

R2: Multi-Scale network exploration. Economists have to understand the national investment and supply chain network to better argue about possible investments, propose tax remodulations, and other types of economic policies. The system must enable network exploration on the different levels of detail on which the information is available. Moreover, the system must implement filter techniques to support the exploration of such a dense network.

R3: Managing queries. During network exploration, analysts regard different aspects of the data by adjusting various filter settings. Often it is necessary to compare and cross-reference the results of such queries. A competent system must thus support the creation, visualization, and comparison of such query profiles.

R4: Identification and tracking of monetary flows. Understanding the monetary flows between sectors and regions could

give a significant advantage in many scenarios, including investment planning, supply chain exploration, and fraud detection. The system should support the users in the identification and tracking of the flows in the transactions network, uncovering paths and patterns pertinent to the application domain.

4. *Hermes* Design and implementation

We designed several prototypes that we used to explore the design space in terms of visualization approaches and interactions methods. During different phases of the project development, the prototypes were inspected and evaluated by our three economists collaborators that supported the requirements definition (see Section 3.3). Our approach was developed using web technology (HTML, CSS, Javascript), as well as frameworks, services, and libraries such as Angular, JQuery, and D3.js. A sketch of the system architecture can be seen in Fig. 1.

4.1. Data model

Hermes data model is represented as a network, whose nodes, called **View Nodes**, represent a productive sector of a specific region (e.g., "Agriculture" in region "Vorarlberg"). Nodes also have attributes, including cash income flow, outcome flow, and growth rate. We obtain the growth rate by calculating the difference of the cashflow (income flow + outcome flow) between the current year (y) and the last year (y-1). Nodes are linked if a financial connection (e.g., investment) exists. We obtain this model from our data sources (see Section 3) as follows: based on the data combination described by Fritz et al. (2003), we link the details about national companies (i.e., expenditures, amount of employees, sectors, locations) to the monetary income and outcome flow of each productive sector of different regions, enabling us to estimate how much money was exchanged between the different sectors and regions. We can, thus, define different *profiles* to filter these nodes. A *profile* is a set of constraints that identifies a subset of nodes. These constraints might include a group of sectors, a group of regions, in- and outflow ranges, as well as a growing rate range.

4.2. Pattern matching algorithm

To create graph patterns, we define and combine specific profiles. These profiles are represented as profile nodes and the user can link them using the Network Board (see Section 4.3.2). Once a pattern has been constructed, we can search for matching subgraphs in the data model. We define $G_t = (V_t, E_t)$ as our data model graph (i.e., the target graph) described in Section 4.1, with V_t the vertex set and E_t the set of ordered pairs of nodes representing the graph edges. We also define $|V_t| = n$. $G_p = (V_p, E_p)$ is the pattern the user constructs with profile nodes (i.e., the pattern graph). We use our pattern matching algorithm to look for a subgraph $G_r = (V_r \subseteq V_t, E_r \subseteq E_t)$, i.e. the result graph. This algorithm conducts the following operations: since G_p nodes represent profiles, which identify a set of nodes in V_t, first (**OP1**) for every profile node $v_p \in V_p$ it extracts the nodes $v_t \in V_t$ that match the constraints defined in v_p and assign them to V_r; (**OP2**) for every node now in V_r, it checks whether its outgoing edges belong to E_r. An edge $e_t = (u_t, v_t) \in E_t$ belongs to E_r if $u_t, v_t \in V_r$ and they were extracted from two different profile nodes in V_p connected by an edge $e_p \in E_p$.

It is immediate to show that our algorithm has a quadratic asymptotic time complexity. *OP1* has to be repeated for each node in V_p, therefore has a complexity of $O(|V_p|)$. We can perform *OP2* in $O(d)$ time for each node, with d the out-degree of a node (i.e., the number of outgoing edges). Therefore, considering that

from V_p we extract the nodes to insert in V_r, the time complexity of the algorithm is $O(|V_r|d)$: $|V_r|$ is at most n, when the queried subgraph includes all the nodes in V_t, and d is at most $n-1$, when v_r shares an edge with all the other vertices in the graph. Therefore, the asymptotic complexity is $O(n^2)$. In practice, we could find that also with very dense graphs our implementation was able to provide the requested subgraph in very short time, thus enabling an interactive exploration (see Section 6). We would like to remark that even if the general pattern matching problem is NP-hard, in our case the problem is greatly simplified because the mapping between the profile nodes and the nodes in the target graph is already known, since is defined by the user.

4.3. Different boards for different tasks

In this section we describe our design of two combinations of different visualizations (i.e., "boards") serving different tasks. In both boards, we used "Color Brewer" (Harrower and Brewer, 2003) to support distinct color pallets for representing different regions. The "multiple boards" approach supports "R2: Multi-Scale network exploration". Both boards are connected by buttons that allow the navigation between each other. It is possible to go from the Overview Board to the Network Board by clicking on the right bottom button "Go To Investigation Mode". To go from the Network Board to the Overview Board, the user can click on the "Go To Overview Mode" button in the bottom center. Active filters from one board are carried to the other to keep the same data investigation context.

4.3.1. Overview board

We designed a fully connected multiple coordinated view board (see Fig. 2) to provide a good overview of region and sector distributions (R1) and to support business contextualization. Views are connected with brushing and linking and filter operations in one view are reflected in all the other views accordingly. Besides providing a good overview, this board is designed to support reasoning about how regions are composed of sectors and how sectors are represented in the different regions. Next, we explain our design choices and the possible interactions for each of the views.

A': Time Slider. The main goal of the time slider is to select data from different years for analysis. The user can drag and drop the slider on any year to load the data of this year.

B': Connection Matrix. We show a preview of connections between sectors for the whole country by using an adjacency matrix. We choose adjacency matrices to visualize a dense network (Ghoniem et al., 2004). Adjacency matrices also have the advantage that they emphasize an additional variable (i.e., in our case the amount of money transacted between two sectors) instead of connections and paths as node-link diagrams do. Moreover when compared to node-link diagrams, adjacency matrices make better use of the space available and they present less visual clutter. Each square represents a connection and the shade of the square encodes the amount of money transacted. Dark areas in the matrix can be spotted immediately which represent big quantities of money transacted. We also implemented a mouse over feature that displays who is the source and target for each matrix connection square, to make precise information about connections between nodes (Keller et al., 2006) and amount of money involved available.

C': Regions Map. The Austria national map including different regions not only works interactively by selecting different region's selectors directly in the map, but also supports the contextualization of physically neighboring regions. Therefore, it also supports presentation purposes.

D': Parallel Coordinates. Each colored line represents a different region while the axes represent the amount of companies
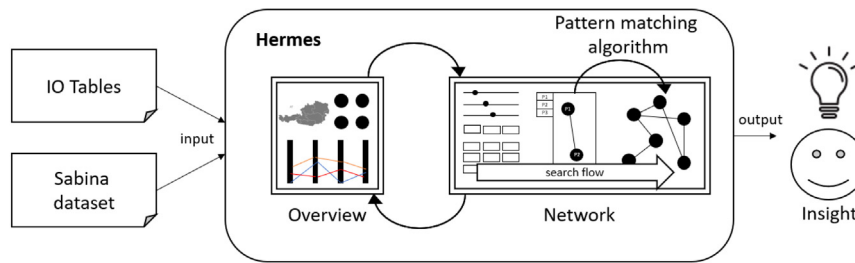
Fig. 1. *Hermes* workflow. Using the two types of input data (left-hand side) *Hermes* presents two connected visualization boards (Overview and Network), which support interactive exploration and insights generations (right-hand side).
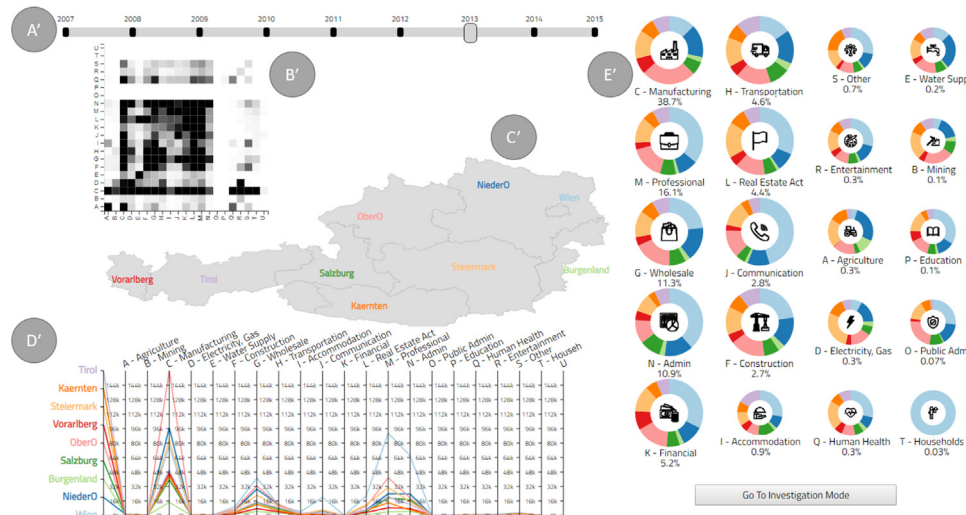

Fig. 2. Screenshot of *Hermes*: Overview Board. (A') Time Slider: allows the selection of different years. (B') Connection Matrix: gives first impressions about how intense the sectors are connected on a national level. (C') Regions Map: shows the color code and the geographical location of the different national regions. (D') Parallel Coordinates: support users to understand how sectors are distributed among regions. (E') Sectors Donut Charts: shows how regions are distributed among sectors.

of each sector. To support different overview tasks, it is also possible to represent other features in the axes such as the total revenue. With this view it is possible to observe how regions are composed by sectors and how sectors are distributed between regions. We opt for using parallel coordinates due to the fact that it supports the representation of multiple dimensions. The axes share a common scale to support comparison.

E': Sectors Donut Charts. A donut chart is essentially a pie chart without the center area. We used the central area for plotting the symbol of each sector. Below each chart, we plot the title of the sector and the percentage that represents the portion of companies within this sector with respect to the all companies in the country (in the selected year). While donut charts should not be used for precise comparisons, we chose them for two reasons (1) to create an area dedicated to easy sector overview and selection (comparable with what view Fig. 2: C') does for regions and (2) to get an immediate impression of the main regions a sector is located in. Also, their aspect ratio allows for the representation of multiple sectors at compact space. In this view, we distinguish between two sizes of donut charts, smaller charts represent sectors that had less than one percent of the national cash flow while bigger charts represents sectors with more than that. This feature facilitates the task of identifying the main sectors of the country. It is possible to select values (sectors and/or regions) from the Overview Board to filter the data that is available in the Network Board, which implements the overview + details on demand mantra (Cockburn et al., 2009).

4.3.2. Network board

During this subsection we refer to view labels from Fig. 3. The Network Board is divided into views A, B, and C to support a three step workflow. In view A, the user creates profiles of sectors and/or regions with specific income and outcome properties as well as growing rates. For example, a profile could be "sectors from Burgenland with a cash income of more than 30 and less than 100 million Euros". In view B, these profiles can then be used as network nodes to construct a network pattern to query the data. The result of a query is displayed in view C. Views D and E present different exploration features that support view C. The workflow presented by the Network Board aims to support "R3: Managing queries" and "R4: Identification and tracking of monetary flows".

A: Profile Creation View. The profile creation view is divided into two parts: (1) the guidance-enriched sliders to filter income, outcome, and growing rates, and (2) the guidance-enriched tiles to select regions and sectors. Inspired by the Scented Widget (Willett et al., 2007), the guidance-enriched sliders are designed to increase user familiarity with the data during the profile creation process. Three sliders are presented: income cash flow, outcome cash flow, and growth rate. The income and outcome cash flow dimensions filter how much money a possible sector or region handles while the growing rate filters the growing (or shrinking) rates of sectors and regions. Growing rates are derived by comparing the analyzed year (y) to the previous year (y-1). The guidance aspects to facilitate informed filtering are described in Section 4.4. Defined profiles are available in view (B) and can be selected as a node to construct a network pattern. Unique profile
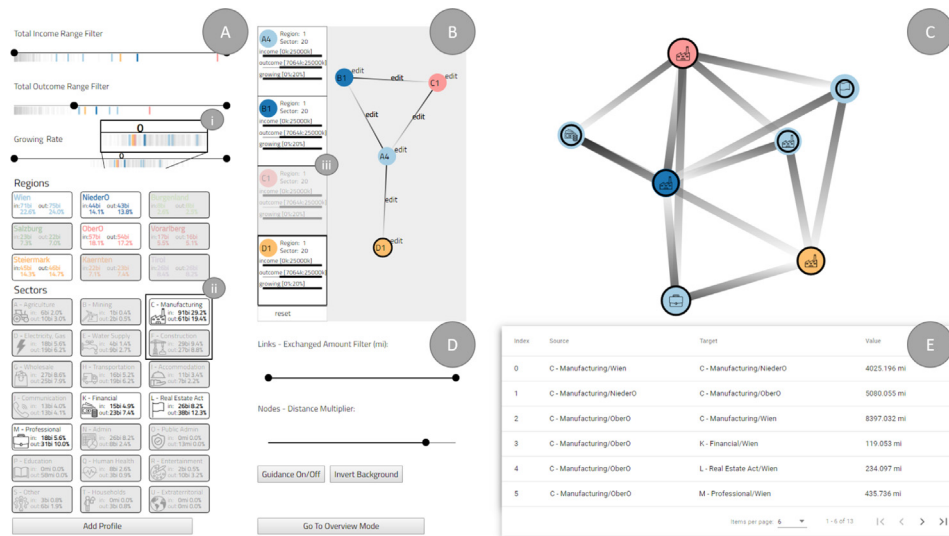
Fig. 3. Screenshot of *Hermes*: Network Board. (A) Profile Creation View: Profiles created in this view are used to query network patterns in view (B). View (A) is divided into two parts: (i) guidance-enriched sliders and (ii) region/sector filter selection. (B) Guidance-Enriched Pattern Search: here the user can link together profiles to create patterns to be matched. The query result is displayed in (C). (D) Control Panel: this panel allows users to regulate some features from view (C). (E) Dynamic Table: this view shows raw transaction data.

IDs are assigned which contain a unique letter + the amount of represented regions/sectors. Profile definitions might be very broad as well as very specific. For instance, a profile which filters only the income dimension to more than 12 billion Euros will match three nodes in the data set described in Section 3.1. On the other hand, if we filter to less than 12 billion Euros, 160 nodes would match the profile.

B: Pattern Search View. In the Guidance-Enriched Pattern Search View, we have (1) a list of created profiles on the left and (2) a drawing canvas on the right. This view allows users to query for network patterns by selecting profile nodes in (1) and drawing connections between them in (2). A visual summary of the created profiles is shown in (1): a circle filled by the regions' color (or gray if multiple regions are involved) and the profile's ID together with small symbol representing the intervals of income, outcome, and growing rate for this profile. The drawing canvas (2) allows the user visually to recognize the query patterns efficiently (Ware, 2012). It supports various profile connection combinations. The canvas also allows for interactively deleting and adding profile nodes and edges as well as changing the direction of edges. Edges might be uni-directional or bi-directional. The generation of profile nodes allows the user to create a compact representation of complex constraints, i.e., multiple sectors and regions with specific characteristics can be combined into one node, which facilitates the construction of network structures to query for and improves clarity of these structures. We present the combination of profiles in (2) as a node-link diagram not only to preserve visual analogy between the constructed pattern and the network data represented in (C), but also to better support path-findings tasks (Okoe et al., 2018). Moreover, compared to other visualization of network data, such as adjacency matrices, node-link diagrams are better suited to communicate and memorize network topology and connectivity (Okoe et al., 2018). These tasks are essential for analyzing economic flow.

C: Node-Link View. We opted for a node-link view instead of adjacency matrix due to two main design advantages: (1) it supports visual matching between the pattern drawn by the user in view B and the query results and (2) it supports better reasoning about indirect paths between two nodes (R2 and R4) (Keller et al., 2006). The compact network pattern created in view (B) is used to query the data set for sub graphs that

match this pattern. The results are then shown in the Node-Link View. However, this view shows a faithful representation of the data set without collapsing multiple sectors into one node, as it is done in view (B) to facilitate pattern construction. Thus, each node in this view represents a specific sector of a specific region and edges represent the yearly money flow between them. The size of nodes is proportional to the amount of money they handle (income & outcome) while the ring close to the outer border of the node, indicates the grow rate by representing the size of the node in the previous year (see Fig. 4). In other words, if the ring of a node is bigger than the node itself, the amount of money handled by this sector was decreasing since last year. On the other hand, if the ring is smaller than the node, this sector has a growing trend (see Fig. 4). The amount of money exchanged between sectors is encoded by edge width. To define this width, we again use a logarithmic scale to avoid extreme visual outliers. Following Holen and van Wijk's (Holten and Van Wijk, 2009) recommendations, we avoided standard arrow representation and, instead, chose a dark-to-light encoding. The money flows from the node at the dark side of the edge to the node at the light side. For bi-direction edges both sides are dark (see leftmost node in Fig. 3). All nodes and edges can be selected for viewing detail information in the Dynamic Table View (see view (E)).

D: Control Panel. The control panel was implemented to facilitate the inspection and refinement of the network presented in view (C). The first slider allows filtering edges by amount of money and the second slider allows re-scaling the network by increasing or decreasing the forces between the nodes. Since edges directions are represented by a dark-to-light gradient, we provide the option to invert the background of the node-link diagram to black, which facilitates the identification of receiver patterns.

E: Dynamic Table. During our problem analysis together with experts, we learned that they prefer to rely on detailed information presented in a spreadsheet-like representation. While tables are not suited to provide a good overview of the network structure, they are good in presenting all the detail information of specific sectors and regions. Thus, we implemented a dynamic table to provide experts with these details on demand. The dynamic table is populated with information about all edges visible in view (C). Selecting edges in this table highlights their representations in view (C).
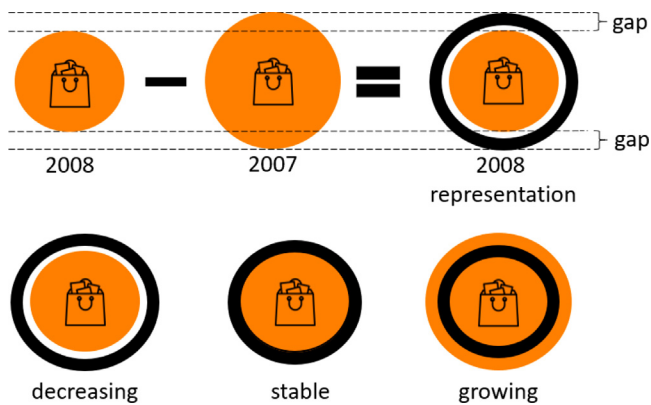
Fig. 4. Visual encoding of the growing rate of a sector. The first row shows the comparison operation and how it is reflected in the visual encoding of a sector node. Using this metaphor, we can visually represent shrinking, stable, and growing sectors (second row).

4.4. Guidance aspects

In this subsection we describe the different degrees and features of guidance that are provided by *Hermes*. Guidance by itself is a computer-assisted process that continuously reduces the knowledge gap that may block the data exploration and analysis. It supports the users to make sense of the data on their own (Ceneda et al., 2017, 2018a, 2019).

A huge number of profile combinations and constructed network patterns would be feasible, but only a small set of them make sense for the data at hand. To avoid frustrating trial and error interaction until a meaningful pattern for querying the data set could be found, we needed to design *Hermes* with features that guide the user in this pattern construction. *Hermes* includes four different features aimed at providing guidance to the user, and thus, facilitating a meaningful and informed usage of the tool. The following indexes "i, ii, iii" refer to Fig. 3 and "iv" refers to Fig. 5.

i: Slider Ticks. Sliders in the Profile Creation View are augmented with ticks colored with respect to the different regions (see Fig. 3: i). One tick represents one sector in one region. Hovering these ticks makes a tool tip window appear which shows full textual information about the sector, region, income, outcome, and growing rate. Ticks are grayed out when they do not match all defined filters. Thus, this visualization of ticks at the point of interaction (i.e., the slider) guides the user in making a meaningful selection of sectors for profile construction, by a priory avoiding constructions that would not lead to any results as these constructs cannot be found in the particular data set. This kind of guidance (i.e., adding information that facilitates the interaction and selection of the data) meets the guidance degree of orienting (Ceneda et al., 2017) which is aimed at building and preserving a user's mental map (Archambault and Purchase, 2012).

ii: Selection Tiles. The tiles for selecting regions and sectors work in accordance with the slider filters, for including or excluding regions and/or sectors from the filtered selection. Options that, in combination with the other filter settings, do not exist in the data are grayed out (see Fig. 3: ii). This, again, helps to avoid the construction of non-existing profiles. Since this feature limits the users interaction possibilities (i.e., grayed-out tiles cannot be selected), it corresponds to the guidance degree of directing, which usually offer the user a pre-selected pool of alternative next steps for a given task (Ceneda et al., 2017).

iii: Guided Network Pattern Construction. While the two previous guidance features help the user to define meaningful profiles, the combination of these profile nodes into a network pattern that can be found in the data, is also a laborious task if performed without guidance. Thus, the Pattern Search View (see Section 4.3.2 (B)) does not allow for the construction of non-existing patterns. This is achieved in a two-step process: First, every time the user selects a profile node that is part of the constructed pattern in the drawing canvas (in Fig. 3: B) in order to add a connection to another profile node, the system runs a query that looks for possible connections from this node in the data set, and disables the selection of all profile nodes that would lead to empty query results. For example, a given "profile C" that would lead to empty query results when connected with "profile A". However, it might be possible to connect it with "profile B". In Fig. 3: iii, the node D1 is selected in the pattern drawing canvas and thus, profile C1 is disabled in the list of available profiles by the system. Again, this feature limits users in their possibilities to construct arbitrary patterns, which makes it directing guidance (Ceneda et al., 2017).

iv: Switching to Bipartite Layout. Another design choice that helps orienting the user (Ceneda et al., 2017), is that the network presentation in the Node-Link View switches to a bipartite layout (see Fig. 5) every time the user constructs a profile combination in the Pattern Search View that connects only two profiles. The automatic adaption of the layout creates an additional awareness in the user that he/she is now investigating a bipartite graph. While this is obvious from the query pattern constructed in the pattern canvas, it can be hard to spot in the Node-Link View which nodes from the query result belong to which profile. Moreover, the layout facilitates the flow analysis between the two groups.

5. Evaluation

To assess and evaluate the effectiveness and limitations of *Hermes*, we conducted a qualitative user study, in the form of an interview with three domain experts (P1, P2, P3) (Isenberg et al., 2013; Kriglstein and Pohl, 2015). We first describe the user study design (see Section 5.1) followed by the participant background (see Section 5.2); afterwards we discuss the tasks and the immediate findings of our study. To complement our study, we interviewed a research expert in guidance (P4), who discussed his/her thoughts about the theoretical and practical implications of the guidance features of *Hermes* (see Section 5.5). Furthermore, we also did examination of the transferability of *Hermes* to Financial Fraud Analysis (FFA) with a domain expert (P5).

5.1. Study design

We conducted our user study as a task-based evaluation: the objective was to evaluate and observe how users operate *Hermes* when asked to solve various tasks. Our study design consists of individual interviews followed by a hands-on session to be carried out by up to two participants at a time. Each interview is structured as follows: (1) participants background interview, (2) introduction of overview and Network Board, (3) hands-on testing, (4) task-based evaluation, and (5) final general feedback discussion. Each session lasted about 90 min. We aimed at making the users familiar with *Hermes* before asking them to perform the task-based evaluation individually and (mostly) unguided. For the whole duration of the session, participants' interaction with *Hermes* was recorded. They were also asked to think-aloud in order to follow their reasoning during the evaluation. The tasks chosen are different for the overview and Network Boards and were designed to evaluate whether *Hermes* could fulfill its design requirements (see Section 3.3). We identified four tasks per board (see Sections 5.3 and 5.4), structured according to the
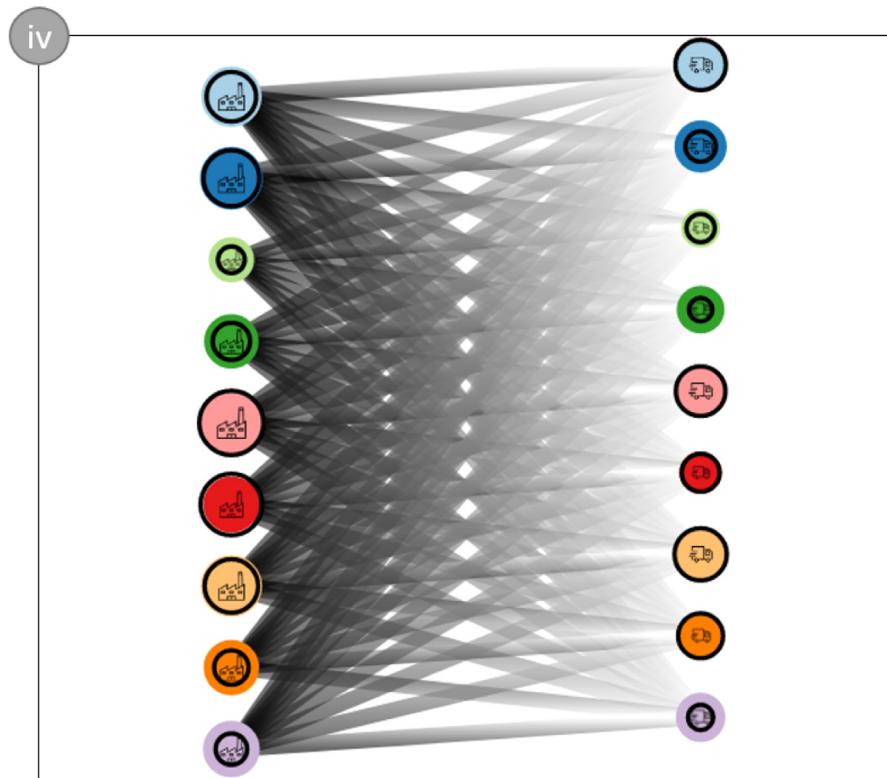
Fig. 5. In this picture, we see a bipartite layout result of two profiles. On the left-hand side, the first profile is representing the manufacturing of each region and, on the right-hand side, the second profile represents the transportation sector of each region. Based on this result, we can investigate how much each manufacturing region spends with each transportation region.

analytical task taxonomy by Andrienko and Andrienko (2006), which highlights differences between elementary and synoptic tasks.

Furthermore, to support insight analysis, we categorize the obtained user insights by following Klein's model of intelligence analysis (Klein, 2013). Our taxonomy includes: **Connection** insights, obtained by combining data from two or more views; **Coincidence** insights, which include relationships between elements or events that are obscured at first, but that the users can perceive as connected; **Curiosity** insights happen when users do not find information when they expect to, and this pushes them to find a reason why. Finally, in order to fully comprehend the impact of our guidance techniques in our study, we implemented in our prototype a counter that would keep track of how many times the system updated the interface to guide the users (see Section 4.4). We called this measurement **Guidance Updates**. The number of insights and guidance updates are displayed in Fig. 6 for each task. During the evaluation session with P1, P2, P3, and P4, we used *Hermes* with the *ESC* dataset (see Section 4.1). The FFA dataset was exceptionally presented to P5, who is domain expert in the field (see Section 5.6).

5.2. Participants' background

P1 and P2: The participating experts (not involved in the design) can all be considered potential *Hermes* target users (see Section 3.1). The two participants are employees from the economic and trade policy department of the "Austrian Federal Economic Chamber" (*Wirtschaftskammer Österreich* or WKO). They analyze the trends of business interests in Austria as well as evaluating and predicting the effects economic policies and regulations throughout the different regions of the country. Both participants use Microsoft Excel (Microsoft, 2018) as well as "Python" and "R"

to explore and present their data. However, they report a lack of visual tools that support their work. Therefore, it is expected that they also lack visual guidance. One participant was familiar with "Tableau", however does not apply it in his work context.

P3: Our third participant is an economist, who works as the global treasury controller within an international company. In contrast to the other two participants, who looked after the productivity of entire regions and multiple companies, P3 is interested in the company's business only. Therefore, P3 is looking for a system that could support investment decisions with information about product supply chain networks. The most used tool by P3 is Microsoft Excel Sheets (Microsoft, 2018) and P3 claims to be familiar with simple visual data representations existing o it.

P4: Aiming for a focused qualitative evaluation of the guidance aspects in *Hermes*, we invited a VA expert researcher with a research focus on guidance. The participant was used to develop VA techniques.

P5: Our fifth participant is a financial fraud expert that works for a national bank. P5 claim being familiar with a good range of visualization techniques and, on her/his work, s/he uses visualizations for presentation and exploration proposes.

5.3. Evaluation of the overview board (P1, P2, P3)

We designed the Overview Tasks (**OT**) to evaluate the first of the two *Hermes*'s boards, which was designed to fulfill requirement R1 (see Section 3.3).

The tasks are the following:

OT1: Which are the biggest sectors in Austria?
OT2: Which regions present more manufacturing companies?
OT3: What is the main region? Why?
OT4: By looking and interacting to the Matrix View (Fig. 2: B'), what type of conclusions could you take?

All the participants concluded the tasks before the specified time, and reported correct answers. Connection insights were the most common insight category during the evaluation of the Overview Board. However, OT4 did not present any connection insight because it is an one view task. Overall, we interpret this result as a positive consequence of presenting a fully interactive multiple coordinate view approach that shows different dimensions of the same data elements. The combination of different selection and filtering features was the main aspect explored during the performance of OT1, OT2, and OT3.

OT3 was especially designed to motivate value thinking by presenting the words "main region" on its specification without describing what was the comparison context encapsulated by the word "main". The participants could define different meanings to the word "main". Still the three participants could explore their hypothesis and reasoning about the different possible outcomes of this interpretation, thus, revealing the flexibility of the tool.

5.4. Evaluation of the network board (P1, P2, P3)

For this board we also conceived four tasks. To evaluate R3 (see Section 3.3) we designed two similar network tasks (NT1 and NT2). The goal is understanding the impact of guidance during the profile construction phase. To do so, we deactivated the guidance features in *Hermes* when evaluating NT1 and reactivated them on NT2. In addition, we asked our participants to tackle two open tasks (NT3 and NT4), which allowed for multiple correct answers. The goal is to observe how many insights and conclusions the participants could draw within the given 5 min time frame. NT3 and NT4 aim at supporting R2 and R4 analysis (see Section 3.3).

NT1: Excluding the region "Oberösterreich", which View Nodes are the sectors with the four highest *income*? How are they connected? The task has to be completed without guidance support.

NT2: Excluding the region "Oberösterreich", which View Nodes are the sectors with the four highest *outcome*? How are they connected? In this task the guidance features are enabled.

NT3: Choose two regions and explore their strongest connections. Discuss your findings.

NT4: Choose a region and search for growing sectors and shrinking sectors. Discuss your findings.

Without guidance support, two out of the three experts were not able to finish NT1. P2 achieved it but past the allotted 5 min mark. Without visual guidance during profile creation, participants were not able to precisely set the income slider to a convenient income range, that led to the inclusion of a large amount of nodes (noise) in each of the created profile nodes. As a consequence, the resulting matched subgraphs did not help comprehending the data and solving the task. On the other hand, the attempts at narrowing the income range failed at including the minimum amount of searched nodes specified by NT1.

With guidance support (NT2), all participants were able to finish the similar task within less than two minutes. Guidance thereby facilitates the determination of a threshold that incorporates data points relevant to the task. By observing the colored ticks under the outcome slider, users could understand how many nodes they were selecting, thus setting a "convenient" slider range. Moreover, since the excluded region is color coded on the Slider Ticks, the participants could exclude samples from the selection without necessarily interacting with selection region filters (see Section 4.3.2), thus speeding up the profile creation process. Tasks NT3 and NT4 are open exploration tasks that were designed to motivate open insights. Both tasks resulted in the most significant amount of insights across all tasks (see Fig. 6).

Overall, the participants provided positive feedback about *Hermes* and complemented different aspects: *"The guidance features*

are really useful" (P1), *"... compared to the technology I use, to perform network exploration tasks, this approach is much faster and appealing"* (P3), or *"...to perform the same tasks with our current approaches is almost impossible"* (P2). P3 also commented positively about the possibility to switch to a bipartite layout: *"...it is good when I search for less groups because the layout gets more organized"*. All participants gave positive feedback on the temporal aspect derived by the growing rate slider. It presented a main role during many of the participants' insights elaboration.

Considering the feedback and the results of the task evaluation, the data appears to support our claims, confirming that *Hermes* fulfills the design requirements. NT1/2 show the capabilities of *Hermes* when handling queries, with the guidance support (R3). The amount of findings in NT3/4 and the feedback gathered suggest that *Hermes* succeeds in enabling multi-scale network exploration (R2) and the tracking of monetary flows and region/sector connection (R4).

5.5. Guidance expert feedback (P4)

The focus of this evaluation was to have P4 to interact with the tool to collect expert feedback about the guidance aspects and not measure insights or answers precision. After performing the complete evaluation, during the post tasks interview, the participant mentioned several positive points. S/He complemented different aspects such as the short interaction response time of the Overview Board and the tasks enabled by the Network Board. S/He highlighted the importance of the guidance features during the profile creation (see Section 4.3.2 view A). Starting with the Slider Ticks (see Fig. 3: i) s/he commented: "*...with the guidance it works perfectly to select the profile that you want to see, otherwise it is complicated to find what you want. It gives you hints about the data that you are selecting*". Concerning "Selection Tiles" (see Fig. 3: ii) and "Guided Network Pattern Construction" (see Fig. 3: iii), P4 considered this guidance aspect as something in between the *orienting degree* and the *directing degree*, more inclined to directing degree (see Section 4.4). Concerning the "Switching to Bipartite Layout" (see Fig. 5: iv), s/he commented: "*it is a good choice that the system supports this automatic decision. It is visually clear for me that this is a better layout for comparison tasks and group understanding*".

When asked about future recommendation for guidance features, a point raised by our participant was the limits of guidance (see Section 6). How much guidance is enough? S/He commented that guidance should always serve a propose, a task. However, when trying to add guidance to many tasks in the same tool, guidance features trend to overcome each other, leading users to confusion. Based on that comment, s/he mentioned that all proposed tasks that were covered by guidance aspects, were not disturbing each other and, s/he would not suggest any further guidance implementation, unless the system is extended to support new tasks. In the same sense, s/he also complemented that the profile creation view does not enforce a specific workflow, avoiding influencing users to limit their profile creation: "*...it is good to avoid steps repetition during exploration tasks*".

Overall, s/he was surprised to see different degrees of guidance in *Hermes*. Her/His favorite features were the guidance implemented in the profile creation and the pattern search view. S/He highlighted several times how much easier and better it was to have guidance features supporting her/him during these tasks.

5.6. Applying Hermes to bank account networks (P5)

We demonstrate the *generalizability* of our approach by applying it to a different domain, i.e., the analysis of financial
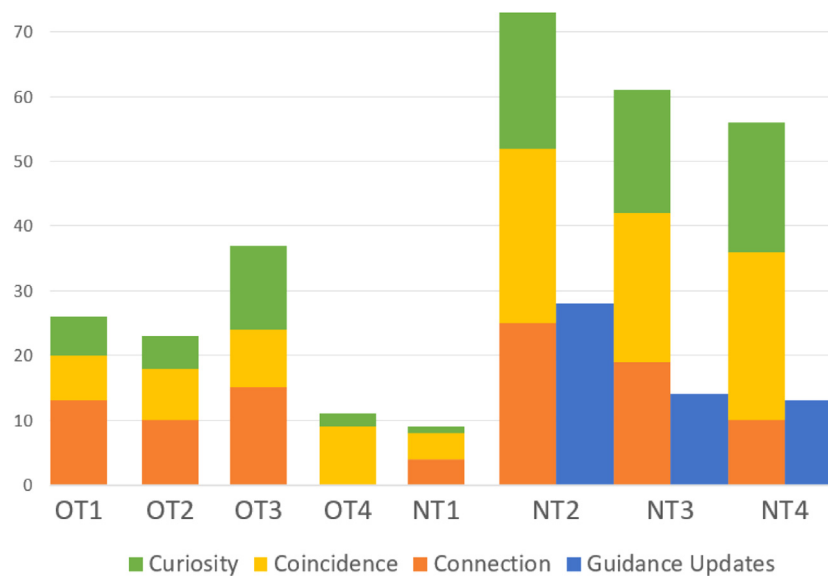
Fig. 6. The overall number of and guidance updates we identified over the different board tasks. We report both the exact numbers in a table and an area chart to easily extract trends from it.

fraud. The FFA data is a real-world dataset for which we render bank accounts as nodes and money transaction between accounts as edges between them. Transaction are ranked according to a metric that measures its how likely it is to be part of a fraudulent behavior calculated by a bank internal automatic algorithm. *Hermes* could be used by fraud investigators for tasks such as identifying terrorists, bank credit exploration, or fraud analysis. For these cases, profiles of bank accounts could be composed of features like age, balance, transaction frequency, account growing rate. With the combination of profile definition and profile-based network pattern queries the investigators would gain a supportive and flexible ally.

We also had the opportunity to present *Hermes* applied to FFA data to a FFA expert (P5), and to collect positive feedback. P5 was able to interact and understand all *Hermes*' features. In particular, the expert appreciated: (1) the potential of flexible profile construction coupled with the network pattern search for detecting networks of fraudsters, (2) the quick pattern search response of *Hermes*, and (3) the guidance aspects during profile creation - which s/he assessed essential for understanding what s/he was selecting and allowing for an informed profile creation.

6. Discussion and future work

In this section we list the limitations of *Hermes* and we discuss some general aspects about guidance that came up during the discussions with the guidance expert. Moreover, we derive directions for future work.

Limited Support of Profile Membership Identification. We provide a feature that changes the layout of the node-link diagram to help the user understand which nodes belong to which profiles (see Section 4.4: iv). By now this feature is limited to bipartite graphs (i.e., including only two profiles). Because this feature was very well received in the evaluation, we plan to incorporate a similar functionality also for queries including more than two profiles. This comes with additional challenges of how to arrange the graph layout when the number of profiles exceeds a certain limit. Radial layouts, for instance, might be implemented to address this problem.

Connections Between Profiles are AND Operations From our requirement analysis we derived the need of analysts to search for patterns where all node types of one profile are connected to

all node types of another profile. Thus, connecting two profiles in *Hermes* queries the data for network structures in which nodes of different profiles with an AND (i.e., all nodes of one profile are connected to all nodes from the other profile). However, new ideas came up in the evaluation and thus, we want to further investigate if connecting two profiles with an OR operation would provide additional benefits to analysts.

Pre-defined Profiles & Network Structures. Another idea stemming from our evaluation with domain experts is to provide pre-defined profiles as well as pre-defined network structures of common queries. For example, experts suggested a profile that involves only elements with a high grow rate or a profile of elements with a high outcome. They also suggested that pre-defined network structures would support an even faster analysis. However, to provide pre-defined profiles and network structures, domain knowledge is required. Moreover, it is not clear how such "ready to use" building blocks influence users in their analysis and if they would limit exploration breadths (Wall et al., 2017).

Performance Considerations. As we described in Section 3.1, we experimented with the ESC dataset, a fairly dense graph. Despite this, thanks to the simplicity of our approach (see Section 4.2), we observed that *Hermes* featured low response times also with more complex queries. During our expert evaluation and internal testing, we tested queries that included stars (a central node connected to some neighbors), cycles, disconnected components, and combinations thereof up to 21 profile nodes (whose constraints identified several nodes in the model graph). We measured the response times, from the pattern extraction to the display of the matched sub-graph, which peaked at 67 ms.

Classification of Guidance. It is not always unambiguous how to classify the provided guidance and even what should be considered as guidance. However, the expert agreed that quite often features might present a mixture of different degrees. For instance, all Selection Tiles that do not exist in the data in combination with the given filter settings are grayed out. This might be considered just a hint for orientation (i.e., orienting guidance), but it also presents users with a list of meaningful alternatives to proceed with, which is considered directing guidance. Sometimes basic visualization features comply with the definition of guidance. For example, in a node-link diagram, encoding the connectivity degree of a node in the node's size to emphasize how well a node is connected might also be considered some sort of

orienting guidance. However, these basic features are common practice in the visualization community and, therefore, are not considered to be guidance. Further research is needed to better distinguish guidance.

7. Conclusion

We developed *Hermes*, a guidance-enriched VA approach to support national economic supply chain analysis on an overview level as well as on a detailed network analysis level. The design and implementation of *Hermes* was inspired by our own research experience and by tight collaboration with economy domain experts, with whom we elaborated a number of design requirements. Our approach consists of two connected boards which provide overview and network views augmented with various guidance aspects to support economy experts with effectively exploring and analyzing complex data.

We designed and evaluated *Hermes* with two real-world data sets and five domain experts (three economists, one FFA expert, and one guidance researcher expert), from whom we obtained positive feedback and further ideas on various aspects of our design. The added value of our approach is: (1) the analysis of insights gained by our evaluators, (2) a positive evaluation of the guidance aspects by a guidance researcher expert, and (3) the demonstration of the transferability of *Hermes* to FFA.

CRediT authorship contribution statement

Roger A. Leite: Term definition, Methodology, Development, Software, Programming, Formal analysis, Application of statistical, Investigation, Conducting a research, Resources, Provision of study materials, Data management, Writing - original draft, Writing - review & editing, Visualization preparation, Supervision, Oversight and leadership responsibility, Project administration, Management and coordination. **Alessio Arleo:** Term definition, Formal analysis, Application of statistical, Investigation, Conducting a research, Methodology, Development, Data management, Writing - original draft, Writing - review & editing. **Johannes Sorger:** Term definition, Methodology, Development, Investigation, Conducting a research, Data management, Writing - original draft, Writing - review & editing. **Theresia Gschwandtner:** Term definition, Methodology, Development, Writing - original draft, Writing - review & editing. **Silvia Miksch:** Term definition, Methodology, Development, Writing - review & editing, Supervision, Oversight and leadership responsibility, Funding acquisition, Acquisition of the financial support.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.visinf.2020.09.006.

References

Aigner, W., Miksch, S., Schumann, H., Tominski, C., 2011. Visualization of Time-Oriented Data. Springer Science & Business Media.

Allen, F., Babus, A., 2009. Networks in finance. In: The network challenge: strategy, profit, and risk in an interlinked world, Vol. 367. Wharton School Publishing Upper Saddle River, NJ.

Andrienko, N., Andrienko, G., 2006. Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach. Springer Berlin Heidelberg, Berlin, Heidelberg.

Archambault, D., Purchase, H.C., 2012. Mental map preservation helps user orientation in dynamic graphs. In: International Symposium on Graph Drawing. Springer, pp. 475–486.

Bigelow, A., Nobre, C., Meyer, M., Lex, A., 2019. Origraph: interactive network wrangling. In: 2019 IEEE Conference on Visual Analytics Science and Technology. VAST, IEEE, pp. 81–92.

Ceneda, D., Gschwandtner, T., May, T., Miksch, S., Schulz, H.-J., Streit, M., Tominski, C., 2017. Characterizing guidance in visual analytics. IEEE Trans. Vis. Comput. Graphics (1), 111–120.

Ceneda, D., Gschwandtner, T., May, T., Miksch, S., Streit, M., Tominski, C., 2018a. Guidance or no guidance? A decision tree can help. In: EuroVis Workshop on Visual Analytics. EuroVA, The Eurographics Association.

Ceneda, D., Gschwandtner, T., Miksch, S., 2019. A review of guidance approaches in visual data analysis: a multifocal perspective. 38, (3), pp. 861–879.

Ceneda, D., Gschwandtner, T., Miksch, S., Tominski, C., 2018b. Guided visual exploration of cyclical patterns in time-series. In: Visualization in Data Science. IEEE Digital Library, Posterpräsentation: Visualization in Data Science (VDS at IEEE VIS 2018), Berlin; 2018-10-22.

Chau, D.H., Faloutsos, C., Tong, H., Hong, J.I., Gallagher, B., Eliassi-Rad, T., 2008. Graphite: A visual query system for large graphs. In: Data Mining Workshops, 2008. ICDMW'08. IEEE International Conference on. IEEE, pp. 963–966.

Cockburn, A., Karlson, A., Bederson, B.B., 2009. A review of overview + detail, zooming, and focus + context interfaces. ACM Comput. Surv. 41 (1), 2.

Didimo, W., Giacchè, F., Montecchiani, F., 2015. Kojaph: visual definition and exploration of patterns in graph databases. In: International Symposium on Graph Drawing and Network Visualization. Springer, pp. 272–278.

Didimo, W., Giamminonni, L., Liotta, G., Montecchiani, F., Pagliuca, D., 2018. A visual analytics system to support tax evasion discovery. Decis. Support Syst. 110, 71–83.

Didimo, W., Liotta, G., Montecchiani, F., Palladino, P., 2011. An advanced network visualization system for financial crime detection. In: Pacific Visualization Symposium. PacificVis, IEEE, pp. 203–210.

Dumas, M., McGuffin, M.J., Lemieux, V.L., 2014. Financevis.net - A visual survey of financial data visualizations. In: Poster Abstracts of IEEE VIS 2014. URL http://financevis.net/, Poster and Extended Abstract.

Fritz, O., Kurzmann, R., Zakarias, G., Streicher, G., et al., 2003. Constructing regional input-output tables for Austria. Austrian Econ. Q. 1, 23–39.

Fujiwara, T., Crnovrsanin, T., Ma, K.-L., 2018. Concise provenance of interactive network analysis. Vis. Inform. 2 (4), 213–224.

Ghoniem, M., Fekete, J.-D., Castagliola, P., 2004. A comparison of the readability of graphs using node-link and matrix-based representations. In: IEEE Symposium on Information Visualization. IEEE, pp. 17–24.

Harrower, M., Brewer, C.A., 2003. ColorBrewer. org: an online tool for selecting colour schemes for maps. Cartogr. J. 40 (1), 27–37.

Holten, D., Van Wijk, J.J., 2009. A user study on visualizing directed edges in graphs. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 2299–2308.

Horvitz, E., 1999. Principles of mixed-initiative user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 159–166.

Huang, M.L., Liang, J., Nguyen, Q.V., 2009. A visualization approach for frauds detection in financial market. In: 13th International Conference Information Visualisation. IEEE, pp. 197–202.

Isenberg, T., Isenberg, P., Chen, J., Sedlmair, M., Möller, T., 2013. A systematic review on the practice of evaluating visualization. IEEE Trans. Vis. Comput. Graphics 19 (12), 2818–2827.

Keller, R., Eckert, C.M., Clarkson, P.J., 2006. Matrices or node-link diagrams: which visual representation is better for visualising connectivity models?. Inf. Vis. 5 (1), 62–76.

Klein, G., 2013. Seeing What Others Don't: The Remarkable Ways We Gain Insights. PublicAffairs.

Ko, S., Cho, I., Afzal, S., Yau, C., Chae, J., Malik, A., Beck, K., Jang, Y., Ribarsky, W., Ebert, D.S., 2016. A survey on visual analysis approaches for financial data. In: Computer Graphics Forum, Vol. 35, No. 3. Wiley Online Library, pp. 599–617.

Koenig, P.-Y., Zaidi, F., Archambault, D., 2010. Interactive searching and visualization of patterns in attributed graphs. In: Proceedings of Graphics Interface 2010. Canadian Information Processing Society, pp. 113–120.

Kriglstein, S., Pohl, M., 2015. Choosing the right sample? Experiences of selecting participants for visualization evaluation. In: Aigner, W., Rosenthal, P., Scheidegger, C. (Eds.), EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization. EuroRV3, The Eurographics Association.

Microsoft, 2018. Excel. http://office.microsoft.com/en-us/excel/. (Accessed 06 October 2018).

Miksch, S., Aigner, W., 2014. A matter of time: Applying a data–users–tasks design triangle to visual analytics of time-oriented data. Comput. Graph. 38, 286–290.

Munzner, T., 2009. A nested model for visualization design and validation. IEEE Trans. Vis. Comput. Graphics 15 (6), 921–928.

Okoe, M., Jianu, R., Kobourov, S.G., 2018. Node-link or adjacency matrices: Old question, new insights. IEEE Trans. Vis. Comput. Graphics.

ÖNACE, 2019. ÖNACE 2008 - struktur. (Accessed 20 July 2019).

Pienta, R., Tamersoy, A., Endert, A., Navathe, S., Tong, H., Chau, D.H., 2016. Visage: Interactive visual graph querying. In: Proceedings of the International Working Conference on Advanced Visual Interfaces. ACM, pp. 272–279.

Shen, Z., Ma, K.-L., Eliassi-Rad, T., 2006. Visual analysis of large heterogeneous social networks by semantic and structural abstraction. IEEE Trans. Vis. Comput. Graphics 12 (6), 1427–1439.

Tekušová, T., Kohlhammer, J., 2008. Visual analysis and exploration of complex corporate shareholder networks. In: Electronic Imaging 2008. International Society for Optics and Photonics.

Timmer, M.P., Dietzenbacher, E., Los, B., Stehrer, R., de Vries, G.J., 2015. An illustrated user guide to the world input–output database: the case of global automotive production. Rev. Int. Econ. 23 (3), 575–605.

Van den Elzen, S., Van Wijk, J.J., 2014. Multivariate network exploration and presentation: From detail to overview via selections and aggregations. IEEE Trans. Vis. Comput. Graphics 20 (12), 2310–2319.

Wall, E., Blaha, L.M., Franklin, L., Endert, A., 2017. Warning, bias may occur: A proposed approach to detecting cognitive bias in interactive visual analytics. In: 2017 IEEE Conference on Visual Analytics Science and Technology. VAST, IEEE, pp. 104–115.

Ware, C., 2012. Information Visualization: Perception for Design. Elsevier.

Wien, W., 2019. Sabina - info - datenbanken. (Accessed 10 June 2019).

Willett, W., Heer, J., Agrawala, M., 2007. Scented widgets: Improving navigation cues with embedded visualizations. IEEE Trans. Vis. Comput. Graphics 13 (6), 1129–1136.