Universiteit Leiden

Master Media Technology

The Co-investment Network of the Dutch Entrepreneurial Ecosystem: A Complex Network Approach

Name:	Alireza Samiei
Student ID:	s2937506
Date:	24-04-2023
Supervisors:	dr. Carolina Mattsson & dr. Frank Takes

Master's Thesis in Media Technology

Leiden Institute of Advanced Computer Science (LIACS) Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Abstract

This thesis explores the evolution of the co-investment network in the Dutch entrepreneurial ecosystem in the period of 2000–2021. More specifically, we use a complex network approach to study structural problems in the distribution and growth of investments in Dutch companies. We use recently available data to build annual networks of financial activities between investors and companies, and co-investors. Based on these networks, we explore several macro, meso, and micro features of these networks. We specifically examine the topological features of co-investment networks and their evolution over time and identify influential Dutch and international investors based on a number of centrality measures. Moreover, we discuss the growth of the co-investment network and analyse the mechanisms behind it. And lastly, we look at communities of investors and their growth, evolution and decline over the years. Our results demonstrate the relevance and importance of studying the financial networks financing startups. Furthermore, our results indicate the prominence of early-stage investors and investments within the co-investment network and shed light on the lack of influential late-stage investors. Our study identifies the absence of the preferential attachment mechanism to explain the growth of the co-investment network. We believe our findings have important implications for a range of stakeholders in the Dutch entrepreneurial ecosystem, ranging from policymakers to investors and founders and emphasise the importance of understating dynamics shaping the Netherlands' startup market.

Acknowledgements

I would like to express my sincerest gratitude to my supervisors, Dr. Carolina Mattsson and Dr. Frank Takes, for their help and support throughout my research. Their constant guidance and insights have enriched my research. I learnt a lot from their constructive criticism and their valuable knowledge, and I appreciate the opportunity to work with them. I also thank my Media Technology professors, their diverse perspectives and insights were instrumental in shaping my academic journey. I would like to thank Cruchbase for providing me with the data used in this thesis. Their cooperation was essential in making this research possible. I am grateful to my family; without their support, my academic journey would not have been possible. Lastly, I would like to express my heartfelt appreciation to Kiran van Hall, for her encouragement and unwavering support.

Contents

Chapter 1

Introduction

The Netherlands is home to one of Europe's most vibrant, dynamic startup markets. With a population of over 17 million, this 22nd smallest country in Europe has one of the highest startup numbers per capita in the world [19]. Several aspects of the Dutch startup ecosystem are promising: In 2021, two out of the ten largest deals by European tech companies were raised by Dutch companies [24]. In the same year, the Netherlands had 24 tech unicorns, meaning startups valued at $1 billion or more. Given that the Netherlands has already given rise to some of the most significant tech companies, including ASML, Bunq, booking.com, Just Eat, and Adyen, to name a few, this is not surprising [24]. However, it is often said that the full potential of the Dutch startup ecosystem is not realised. One of the most prominent setbacks of the Dutch ecosystem is the lack of capital and structural issues regarding how investments are distributed across the ecosystem [36, 37]. Numerous issues in the Dutch startup ecosystem are related to how funding for startups has evolved over the years, from a lack of funding for university spin-off startups to the untapped potential of deep tech startups, which create and commercialise technological innovations based on engineering expertise and scientific research and need significant investment [35, 37]. The distribution of funding in an ecosystem results from a complex interplay between the components present in the ecosystem. Therefore, to understand the funding of Dutch startups, it is crucial to gain insight into the activities involving investors and startups that make up the components and dynamics of the ecosystem.

Before going any further, we must first clarify what we understand as the ecosystem's dynamics and components. Entrepreneurial activities range from human capital embodied in innovators or institutions, such as universities, to the financial industry and regulatory frameworks and the overall behaviour of the economy at both macro and micro levels. Even though the study of entrepreneurship goes all the way back to the beginning of the 20th century, the focus on the entrepreneur as a systemic concept is relatively new [1]. This shift of focus gave birth to the concept of entrepreneurial ecosystems (EE). Scholars have proposed several definitions to pinpoint the meaning and idea behind the EE concept exactly. Here we mention one that captures the essence of this idea; an EE is:" a set of interconnected entrepreneurial actors (both potential and existing), entrepreneurial organisations (e.g., firms, venture capitalists, business angels, banks), institutions (universities, public sector agencies, financial bodies) and entrepreneurial processes (e.g. the business birth rate, numbers of high-growth firms, levels of blockbuster entrepreneurship, number of serial entrepreneurs, degree of sellout mentality within firms and levels of entrepreneurial ambition) which formally and informally coalesce to connect, mediate and govern the performance within the local entrepreneurial environment" [1, 29]. By considering this definition, we can conclude that an EE is a complex, dynamic, adaptive system that uses and explores resources in a given environment. The EE's concept suggests a framework that centres around a system or network comprised of interconnected components and their complex interactions [1]. Therefore, the possibility exists to conceptualise a network approach to represent EEs. For our research interest, which falls under the umbrella of economic networks, we examine co-investment networks in the Dutch ecosystem to gain insight into the financing of Dutch startups. As Arthur mentioned [2], networks arise in various ways in an economy, and multiple aspects of them are of interest: "how their structure of interaction or topology makes a difference; how markets self-organise within them; how risk is transmitted; how events propagate; how they influence power structures." Within this framework, we can study various types of economic networks, like the co-investment network we are interested in this thesis. We can study features such as the topology of the network in question, power laws and the question of formation in the economy [2].

Based on real-world problems regarding financing startups in the Netherlands, we will focus on the financial network supporting the Dutch EE. Much has been done on the topic of investors who join together and share their investments in literature, which we will go through in the next section. Although there is much knowledge to be gained from an entrepreneurial finance and managerial study approach to an EE's problems, this thesis will look at the co-investment of investors in the Netherlands as a complex network. We will study the co-investment network with regard to its topological macro structure, look at influential investors on a micro-scale and explore the question of community formation and the evolution of the network's meso structures across years. Hence the main research questions of this thesis are:

- What are the topological features of the Dutch co-investment network, and how have these aspects evolved over time?
- Who are the most influential nodes able to transmit information, influence trends of investments and act as bridges across disconnected investors?
- What explains the mechanisms behind the growth of the Dutch co-investment network? Can we find evidence of the preferential attachment phenomena?
- What is the community structure of the Dutch co-investment network? How did these communities evolve over time?

The remainder of the thesis will proceed as follows. In Chapter 2, we begin by reviewing related literature and providing the context of this study. Chapter 3 describes the data, and a brief comment on data limitations and scope is provided. In Chapter 4, we outline the ecosystem network and the one-mode co-investment network, the procedure used to construct them and review their overall structure. Moreover, we describe methods and techniques used to investigate our research questions. In Chapter 5, we paint a general picture of investments in the Dutch ecosystem between 2000–2021 by providing an exploratory analysis of the data. We then move to analyse cumulative networks consisting of Dutch and international investors and look at the growth of this network within the time window of this study. We analyze cumulative co-investment networks using a variety of static network metrics and identify influential investors utilising a range of centrality techniques. Moreover, we will study preferential attachment in the co-investment network and detect its absence or presence. Additionally, the process of community formation will be explored. Lastly, in Chapter 6, we discuss the obtained results and conclude this thesis, and we will provide suggestions for future works.

Chapter 2

Background and Related Work

It is a common saying within the world of financing startups that investors are operating on the basis of a power-law distribution [28]. Although the objectivity of observing power laws in venture investment is a matter of discussion, what is usually referred to as power law within this field means that most of the ventures do not yield any return. Because most startups fail, investors seek out the few ventures that can compensate for all of their failed investments; in other words, they seek out the outliers, startups whose performance equals or outperforms the rest of their investments. Failing to find those few very profitable investments could spell major financial trouble for investors. Hence, minimising this risk is one of the biggest challenges investors face. For numerous reasons that we will consider here, investors sometimes co-invest together. As previously discussed, there are different aspects within the framework of EE that can be interesting and important for research in the Dutch ecosystem; however, here, we focus primarily on providing context on why and how co-investments occur. In this section, we will briefly cover related literature on why co-investments occur and go through a number of studies that analysed co-investment networks more specifically.

Several reasons have been given for why investors prefer co-investing. Firstly, if investors have shared investments before, this will reduce the partner-specific risk and helps the syndicate to have a higher chance of performance [5, 21]. A syndicate, in this context, refers to a group of investors who share their pool of resources to accomplish the goal of growing and eventually selling their ventures [26]. Furthermore, by forming syndicates, VC firms gain better access to the deal flow of other VCs. The rate at which investors get business and investment offers is simply referred to as deal flow. Additionally, the portfolio companies of well-networked VCs have a better chance of successfully exiting the ecosystem [21]. Hence syndication among investors is widely regarded as a measure to have better access to new opportunities, and overall it improves the chance of success of both investors and startups. Moreover, one study showed that due to the high-risk condition that younger VCs face when investing, they tend to enter syndicates with other firms to reduce the risk factors more severe for them than for more established investors [22]. According to the same study, VCs benefit from co-investing in general because they can provide their portfolio companies with a larger pool of managerial advice and expertise. It is important to note that startups look for investments not only for the capital injected into their firms but also for the industry-specific expertise, advice, and even potential customer pools that investors can provide. In [20], Hochberg *et al.* examined the benefits of firms co-investing together in greater depth. They argue that the primary concern of firms in finding partners is a desire for resource accumulation and not, as it is more traditionally argued, a preference for similarity. Interestingly, their study emphasises that investors form ties based on the principle of complementarity, meaning a beneficial association exists between a highly experienced firm with access and investment scope and a capital-rich firm with few of these resources.

As mentioned earlier, one of the primary motives for forming syndicates is to reduce the associated risks of investments. This risk is even more pronounced when investors show interest in getting involved in foreign and especially emerging markets. In [25], Khavul and Deeds studied how the initial ties are formed between domestic and international investors in an emerging market. They found that sharing an interest or expertise in a similar industry is essential for foreign investors trying to select other domestic VCs to invest with. The same behaviour is also present in investment partner selection between foreign investors creating syndicates to invest in an emerging market. More importantly, they established that when foreign VCs made their initial co-investments, they continued to make investments and engaged in new syndicates with new partners. It must be said that much evidence has been found on the positive effects of co-investments. Still, there are also risks associated with syndication [5], like issues regarding decision-making and monitoring between VCs and company founders going through the process of IPO, which is offering their shares to the stock market [14]. Besides the costs related to decision-making and potential conflicts of interest between founders and investors, problems regarding friction in the VC syndicates are also notable. In all of the research mentioned above, much attention has been given to the co-investment network of VC firms. However, several studies focused on other types of investors, such as individuals independently funding startups, i.e., angel investors. In [39], it was shown that co-investment networks also exist among angel investors. More importantly, angels with a higher rate of successful investments have better opportunities to co-invest in more rounds, either as a round leader or as a participant. This study again emphasises the importance of successful investors for startups since ventures that receive investments from successful angels have a better chance of raising later-stage funding rounds, especially from VCs. For example, when raising funds for the next rounds of investment, a syndicate usually adds new investors to its already existing set of members [30]. In the mentioned scenario, the amount and level of access and knowledge of how the startup operates differ between new and old investors.

Lastly, a number of studies specifically applied a complex network analysis approach to the behaviour of VC firms. In [40], Zhang *et al.* studied VCs' investment and co-investment networks in China from 1994 to 2014. They analysed several aspects of temporal networks of VC investment, ranging from the expansion of the VC market to studying the changes in the connected components during a boom period. In another study conducted on the Chinese VC market [27], the performance and importance of VCs between 2009 and 2020 were assessed using a complex network approach. They found that VCs that had good investment results were more inclined to establish co-investment relationships with other firms.

In conclusion, co-investment and syndication are beneficial for various reasons to both investors and companies. Co-investment facilitates investment by mitigating risk for investors and pooling resources between firms, among other factors. Therefore, to understand the financing of the Dutch EE, gaining insight into these co-investment networks present in the Dutch EE is crucial. Hence, this will be the focus of the following sections of this thesis by pursuing a network-based approach.

Chapter 3

Data

This chapter provides an overview of the data used in this thesis. First, we describe the structure of the dataset and give a set of definitions of terms used in the various parts of the dataset. Secondly, we provide a number of remarks on the criteria used in the data selection procedure and mention the important limits and scope of the data used in this thesis.

3.1 Dataset

The data used for this research was acquired from Crunchbase on 26 October 2022. Crunchbase is a San Francisco-based company founded in 2007 that provides business data on public and private companies. As the company's website describes: "Crunchbase has best-in-class live data powered by our unique community of contributors, partners, and in-house data experts. Data is enriched, cleansed, verified, and updated daily to ensure our customers have the latest information on private companies" [12]. Although Crunchbase's database is not intended for the specific usage of economics research, as it is argued in [6, 13] figures produced by other reliable sources were similar to Crunchbase's statistics on venture capital investments. Because of the interconnected nature of the data points in this dataset, Crunchbase's data provides valuable data and insights. The process of creating a dataset that captures the dynamic nature of the Dutch startup ecosystem is discussed later in this chapter.

As for gaining access to the data, an application was sent to Crunchbase's Academic Research Access Program to access the company's dataset.After the access was granted, the Daily CSV Export option was used to gain access to Crunchbase's data. The dataset contains seventeen .csv files; however, for this thesis, only five datasets were used:

- Organizations.csv - organization profiles available on the Crunchbase platform.

- Investors.csv - active investors, including both organisations and people.
- Investments.csv - all investments made by investors.
- Funding Rounds.csv - details for each funding round in the dataset.
- Acquisitions.csv - list of all acquisitions available on the Crunchbase platform.
- IPOs.csv - detail for each IPO in the dataset.

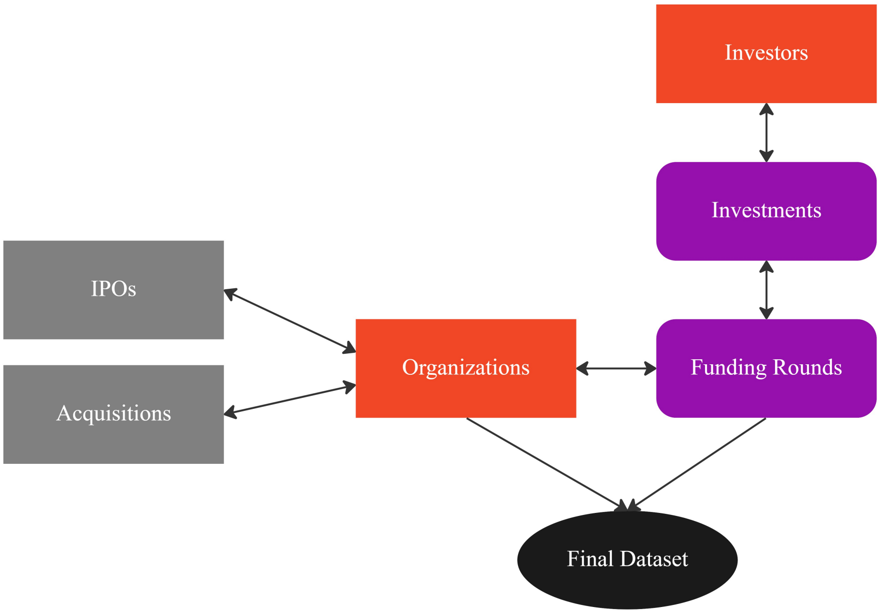

Figure 3.1: The procedure utilised to generate the dataset used in this study

To build the two-mode network of the ecosystem, a network of investments in Dutch startups categorised by investors and startups, and the network of co-investment in the Dutch ecosystem, a finalised dataset was created using the cross-linked information on organisations that received investment in various funding rounds from investors. Figure 3.1 illustrates the process. Firstly, the organisation dataset was filtered by only including organisations with headquarters in the Netherlands, which implies that the scope of this study is limited to Dutch companies; nevertheless, unlike startups, investors can be also international. Hence, our data contains all investments made in Netherlands-based domestic and foreign startups. Only organisations with information on the number of funding rounds were included; out of 50,168 organisations in the Netherlands, only 3,006 had data on the number of funding rounds and were included. This step was necessary to only include organisations with available data on their funding status. The high number of startups with no

Entry	Description of the field	% of missing Values
name	Name of the organisation	0.00
org_city	Where the organisation is headquartered, city	0.00
employee_count	Total number of employees of organisation	0.00
org_status	Operating status of organization(e.g. Active)	0.00
org_num_funding_rounds	The total number of funding rounds	0.00
org_founded_on	The date the organization was founded	1.90
went_public_on	The date when the organization went public	97.21
acquired_on	The date of the acquisition	87.39
category_groups_list	Superset of industries (e.g. Software, Mobile, Health Care)	1.51
investment_type	Type of funding round(e.g. Seed, Series A, Private Equity, Debt Financing)	0.00
raised_amount_usd	Amount of money raised in the funding round (USD)	31.35
investor_count	Total number of investors in a funding round	0.00
post_money_valuation_usd	Valuation of a company after a funding round(USD)	96.44
announced_on_year	The date when investment is announced(Year)	0.00
is_lead_investor	Whether an investor led or organised the investment	52.53
investor_name	The name of the investor	9.64
investor_type	Describes the type of investor this organisation or person is (e.g. Angel, Fund of Funds, Venture Capital)	9.64
country_code	Where the investor is headquartered, country	8.81

Table 3.1: All entries in the finalised dataset used for this study, their descriptions and percentage of missing values

information on their funding status is due to the fact that many startups go out of business without receiving any funding, or simply because the record of their funding was not added to our dataset. Since the scope of this study is focused on co-investments, including startups without any relation to investors was not useful. Another filter was set only to include organisations with the 'company' tag since out of 3,006 remaining organisations, 65 had mixed roles such as 'investor, company', 'investor', 'investor, company, school' and 'company, investor'. Considering the small number of the mentioned organisations compared to the rest of the entries and the fact that the true nature of their financial activity was not verifiable, we decided to filter them out. After filtering out duplicates, 2,466 organizations remained in the dataset. Moreover, the mentioned mixed-tagged companies would create problems in building the two-mode network, which will be discussed later. Using the cross-linked feature of datasets described before, we identified useful data from each dataset. We merged all datasets together by linking the universally unique identifier (uuid) of entities we were interested in and created the final dataset. The merged dataset became available, containing data on each round of funding, including all the necessary data on each round of investment in companies by investors. An overview of all rows included in the dataset is presented in Figure 3.1. Moreover, an example from the final dataset is provided in Table 3.2. An overview of various aspects of this dataset will be explored in Chapter 5.

3.1.1 Cumulative Scope

The dataset used in this research was built cumulatively, meaning that it starts in the year 2000, including all investments that occurred in that year. As we move to next year, all the investments in 2001 were added to the already existing dataset. Therefore, the cumulative dataset for the year 2021 includes all investments from 2000 to 2021. We must mention that when a startup exits in the ecosystem, whether by acquisition, IPO or closure, it still remains in the dataset. In general, this

name	BUX
org_city	Amsterdam
employee_count	101-250
org_status	operating
org_num_funding_rounds	8
org_founded_on	2014-10-01
went_public_on	NaN
acquired_on	NaN
category_groups_list	Financial Services, Lending and Investments
investment_type	series_c
raised_amount_usd	2,460,327
investor_count	3.0
post_money_valuation_usd	NaN
announced_on_year	2017-10-22
is_lead_investor	True
investor_name	HV Capital
investor_type	venture_capital
country_code	DEU

Table 3.2: An example of a funding round which includes a startups and investor

is done because we are more interested in analysing various aspects of co-investments rather than forecasting the success of business ventures, which is a more studied body of literature in this field.

3.1.2 Remarks on Limitations of the Data

In the Crunchbase dataset, not all data points have complete information. The proportion of missing or partially missing values from each data point is shown in Table 3.1. Another notable limitation of this dataset is the time the frequency of records added to the dataset. The Crunchbase dataset was created in May 2007; entries before this date have been added later to the dataset. Following the exploratory analysis and, later on, the network analysis provided in Chapter 5, it is important to have the limitations of the scope and depth of this dataset in mind.

Chapter 4

Methods

In this section, we outline the methods and procedures utilised to construct and investigate the the ecosystem and co-investment networks. This chapter aims to provide a set of definitions, explanations and examples to describe methods used in the research process. We go through networks used in this thesis and then describe static network metrics and their relevance to answering research questions. Moreover, we discuss centrality measures and the community detection method used in this research.

4.1 Networks

Studying the networks used in this research begins with understanding their components and their interactions. Given that we used two sets of networks with different properties, in this section, we review their structure, how they were created using the finalised dataset, and mention a number of the network's characteristics.

4.1.1 The Two-Mode Ecosystem Network

A two-mode network consists of two distinct sets of nodes such that links only connect nodes from the opposite set [4]. Another key consideration when examining network links is whether they are weighted or unweighted. In weighted networks, each link has a unique weight [4]. The first network we investigate is the two-mode investors and startup network, which we call the ecosystem network. The ecosystem network consists of investor and startup node sets. Each link is a funding round, establishing a connection between investors and startups. As a result, links signify a financial association, and the network can be viewed as a financial network. We decided not to include weights for links in this setup. Including weights for each link would be beneficial; however, approximately 30% of the amount of money raised for these rounds was absent from the original dataset. Finally, the co-investment network is the primary focus of this thesis, and as we will see, we emphasise the number of prior investments as the defining factor for connections between investors rather than the amount of money involved in each round.

4.1.2 Constructing the Two-Mode Ecosystem Network

Based on the finalised form of the cumulative datasets, annual two-mode networks containing funding rounds in the Dutch startup ecosystem were created, resulting in 21 networks corresponding to the years of the time window in this study. Each network contains investments from that year and all investments that happened beforehand. Furthermore, the ecosystem networks allows the existence of parallel links. Parallel links or multi-links are two or more links that connect the same pair of nodes in a network. Since multiple examples of repeated investment existed in the data, we used parallel links to represent multiple rounds of investments. An important point to mention is that although networks are constructed annually, startups that leave the ecosystem are not filtered out. This is due to the fact that startups leave the ecosystem based on three different scenarios: IPOs, acquisition and closure. There is data available on IPOs and acquisition; however, the biggest reason for startups to leave the ecosystem is to go out of business at some point, and there is no data available on the date when the closure happened. After the cumulative two-mode networks were built, they were examined based on a number of network metrics later discussed in this chapter. A simple schema of the ecosystem network is depicted in Figure 4.1, along with all of the attributes used for nodes and links.

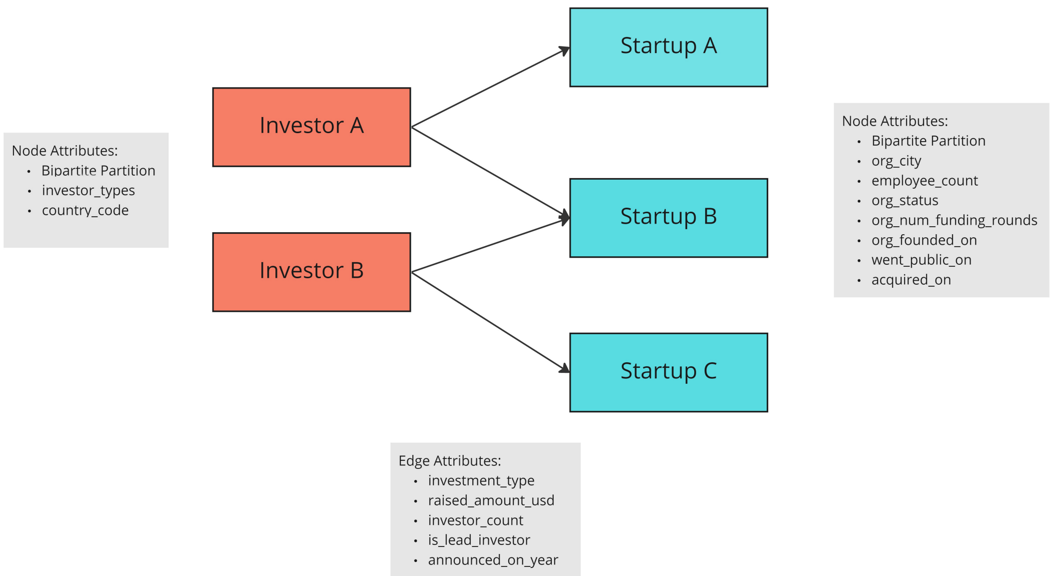

Figure 4.1: A simplified representation of the two-mode network, plus node and edge attributes

4.1.3 The One-Mode Co-investment Network

It is possible to generate two one-mode projections for each two-mode network with two distinctive node sets. A projection of a two-mode network is the transformation of each set of its nodes into a one-mode network. Links between nodes in the projected network exist if they share a common neighbour in the two-mode network's opposite node set [4]. In this thesis, we only examine the projection of the investor's node set. The one-mode investor's network contains prior co-investments by investors in the startups.

4.1.4 One-mode Projection of Investor Node Set

A weighted one-mode projection containing only nodes from the investor's partition was created for each two-mode network corresponding to the time window of this study. These projected networks are weighted, with weights denoting the number of shared investments in the two-mode ecosystem network. Figure 4.2, depicts the process of projecting the two-mode network into a one-mode network.

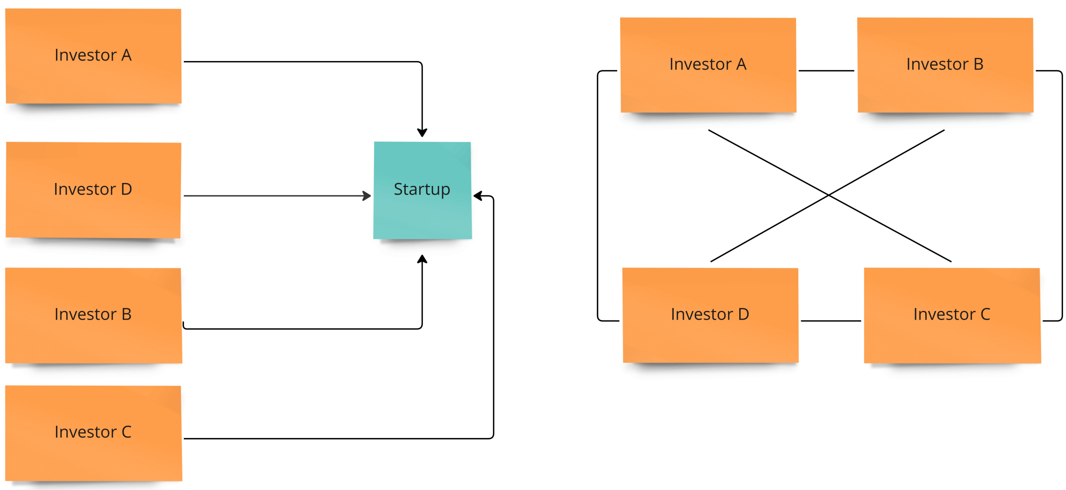

Figure 4.2: Projection of the investor's node set based on co-investment

4.2 Measures for Analysing the Structure of Networks

This section explains various methods to analyse the ecosystem and co-investment networks. These methods are used later in Chapter 5 to answer the research questions of this thesis.

4.2.1 Degree and Average Degree

The degree of a node is the number of all links connected to the node [23]. For weighted networks, we consider the weight of links while calculating the degree. A useful way of putting the degree of nodes in a network into perspective is by looking at the average degree of the network. The average degree is the division of the number of links by the number of nodes present in the network. In the ecosystem network, an investor with a high degree score shows a high number of investments; likewise, a startup with a high degree score is a company with multiple received investments. As for the co-investment network, an investor with a high degree has a high number of shared investments with other investors.

4.2.2 Density

The density of a network is the proportion of present links to the total number of possible links in the network [23]. A low density indicates the sparsity of the network, and a high density score indicates a cohesive network. If the ecosystem scores high on density, it means that a high number of investments have been made in the network. Likewise, a highly dense co-investment network means a large number of investors sharing prior co-investments together.

4.2.3 Degree Distribution

A network's degree distribution represents the proportionate occurrences of nodes with varying degrees [23]. Across different networks, various degree distributions are observable. For example, a regular network, where all nodes have the same degree, exhibits a wildly different degree distribution compared to a network with heterogeneous degrees, such as a random network [23]. Most real-world networks follow a heavy-tailed distribution when plotting their degree distribution. In simple terms, these heavy-tailed distributions show a handful of hubs with a large number of links and many nodes with a small number of links [4]. Later, we will explore the degree distribution for both ecosystem and co-investment networks.

4.2.4 Degree Assortativity Coefficient

Assortativity is the tendency of nodes to form connections to similar nodes rather than dissimilar nodes. More specifically, assortativity by degree simply means the tendency of nodes to connect to nodes with a similar degree. For calculating the degree assortativity coefficient, we used methods proposed by Newman [31] and Foster *et al.* [15]. The value of this measure is between -1 and 1. For the co-investment network, 1 indicates the presence of a tendency of investors with a high number of co-investments to connect to other high-degree investors. In contrast, the degree assortativity coefficient of -1 indicates perfect disassortativity, meaning there is a tendency for active investors to share investments with investors with a small number of prior co-investments. We only measured the degree assortativity coefficient of the co-investment network.

4.2.5 Connected Components

First, we need to explore two concepts of connectedness and components to understand what connected components are. Two nodes are connected if a path exists between them. A path is a sequence of links between two nodes. Based on this notion, a network is called connected if every node is connected to every other in the network via a path. On the contrary, if one pair of nodes exists in a given network without a path between them, the network is disconnected. Based on this notion of connectedness, a connected component is a subset of the network that all pairs of nodes in that subset are connected by a path [4]. Many networks, such as the co-investment network, are composed of multiple separate components and hence are not fully connected. In real-world and random networks, when networks start to grow, we usually observe that a significant proportion of nodes form a giant component, and a lot of smaller components are disconnected from the main giant component. If a component of the network contains a significant proportion of nodes in the network, we call that component a giant component (GC). In both ecosystem and co-investment networks, we will study the percentage of nodes and edges in the GC. These GCs are the core of both networks, and studying them more extensively using community detection algorithms gives insight into the structure of the networks in question. Moreover, a number of network measures, such as diameter and average path lengths, are often measured in the GC.

4.2.6 Diameter and Average Shortest Path Length

Here we define important concepts regarding paths and distances in networks. We call the path length the total number of links in a given path [4]. Furthermore, the shortest path is defined as a path between two nodes with the shortest distance, meaning the smallest number of links. Based on the notion of the shortest path, we can now look at network diameter, which is the maximum shortest path in the network. Although diameter is a useful method, we used it in combination with the average shortest path length because it only measures the distance between one pair of nodes at the distribution's extreme end [23]. The average shortest path length is the average number of the shortest paths for all node pairs in a network. These two measures help to understand the number of steps required to find prior ties considering shared investments in the co-investment network and give useful information regarding the network's topology.

4.2.7 Average Clustering Coefficient and Transitivity

Measures discussed here capture the cohesiveness of networks; therefore, they are useful for understanding clustering in networks. Both average clustering coefficient and transitivity are based on the notion of triangles or three-clique in networks and are probabilistic methods. A triangle is a set of three nodes in which each node has a link to the other two nodes [4]. The first method to assess clustering is the average clustering coefficient. The local clustering coefficient is measured by the local density of connections in a node's neighbourhood. The network's average clustering coefficient is then calculated by taking the mean of all local clustering coefficients [34]. Basically, the average clustering coefficient indicates the probability of two neighbours of a randomly selected investor, sharing prior co-investments with each other. The next method is transitivity or the global clustering coefficient, which is defined as the fraction of all possible triangles present in a given network [4]. This measure is calculated as three times the number of triangles over the number of connected triplets. These connected triplets are, in the co-investment network, a set of three investors, such that one investor shared two investments with its neighbours. Still, those neighbours are not connected together.

4.3 Centrality Measures

Measures that were discussed in the previous section are useful to gain a better understanding of the macrostructure of a network [23]. In this section, we explain centrality measures used to identify influential investors with the ability to transmit information, influence trends of investments and act as bridges across the network.

4.3.1 Degree Centrality

First and foremost is the degree centrality, which measures the number of nodes a node is connected to in the network [18]. In social networks, degree centrality is a measure of popularity. In the co-investment network, a high degree centrality score indicates investors who share the most number of co-investments with other investors. This centrality measure is used to further investigate the type and nationality of disproportionately over-represented investors in the network. The degree centrality method we used is unweighted.

4.3.2 Betweenness Centrality

This centrality measure is defined as the sum of the fraction of all pair's shortest paths that pass through a node. The algorithm used to assign betweenness scores to nodes is based on the implementation of Brandes [10, 11] as it was first proposed by Freeman [16]. In its simplest form, Freeman's betweenness centrality indicates a node's importance in connecting other nodes in the network; therefore, this centrality measure captures how important a node is in the transmission of information or resources between other nodes in the network [7]. An investor with high betweenness centrality acts as a broker or bridge, connecting disconnected investors. It can be the case that an investor with a low degree centrality has a high betweenness centrality; meaning that it did not have a lot of co-investments in the network. Still, it connected different groups of investors by sharing a small number of investments with other players who did not have prior ties between themselves. On the contrary, there exist investors with a high degree centrality who also scored high betweenness scores since they connected otherwise disconnected investors, acted as intermediaries, and were in a unique position to coordinate other groups in the network. The betweenness centrality algorithm used in this thesis is weighted.

4.3.3 Eigenvector Centrality

Although the number of co-investments is a good indicator of an investor's influence, with whom those investments are shared is also important; hence eigenvector centrality is also used supplementary to centrality measures mentioned before. This centrality measure scores nodes based on the centrality of their neighbours [18]. As proposed by Bonacich [9], the eigenvector centrality of a node is proportional to its neighbour's centrality. Moreover, this centrality measure is also usable for the weighted networks, as is done in this thesis [7]. Essentially, sharing investments with other active investors increases an investor's eigenvector centrality. It is possible that an investor indicates both a high degree centrality and eigenvector centrality, indicating the high number of co-investments and sharing those investments with other active investors. Still, a high eigenvector centrality is also possible when an investor is involved in funding rounds with other prominent investors. Nevertheless, the investor themselves did not invest as often as their neighbours.

4.4 Preferential Attachment

As for answering one of our research questions regarding the growth mechanism of the co-investment network, we study preferential attachment (PA) in the co-investment network. It was suggested by Barabási [4] that the scale-free property we see in many different networks is due to the coexistence of network growth and PA. Moreover, in [3], it was proposed that the mentioned scale-free properties are due to the addition of new nodes to the network and, more importantly, the fact that these new nodes connect preferentially to well-connected nodes. In the context of the co-investment network, if PA exists, it would result in a disproportional gain of co-investments by highly active investors. To detect PA in the co-investment network, we use an experimental method suggested by Barabási [4]. We create maps of networks across the period of four years and measure how many new co-investments each investor gained during each year. We call this gain of degree, cumulative growth. By plotting cumulative growth on a double logarithmic scale, we detect if cumulative growth indicates the presence of PA, marked by a quadratic function. It can also be the case that cumulative growth is indicated by following a linear or sublinear regime, revealing the absence of PA.

4.5 Community Detection

A community is a group of nodes that are more likely to connect to each other compared to nodes from other groups [4]. As for this thesis's research question regarding co-investment community identification and their evolution over time, we are interested in finding communities of investors that are clustered together. Studying communities and their internal structure gives us a better understanding of the internal interactions of investors and helps us to get a better sense of how these communities were formed and continue to evolve. To achieve this task, the Louvain Method was employed, as explained in Subsection 4.5.1.

4.5.1 Modularity Optimization using the Louvain Method

The Louvain method is based on heuristic modularity optimisation [8]. Modularity, as defined by Newman and Girvan [33], measures " the fraction of the edges in the network that connect vertices of the same type (i.e., within-community edges) minus the expected value of the same quantity in a network with the same community divisions but random connections between the vertices." If modularity indicates 0, it means that the number of internal links in a community is no different than in a randomised network. Positive values of modularity mean the presence of links between similar nodes is visible, and negative modularity scores indicate that fewer number of these links are present in the network compared to a random network [32]. Based on this notion of modularity, the Louvain method investigates each node's modularity score by first calculating the score in its own community, inserting every node in its neighbouring communities and repeating the process until the modularity score of a node in a community is maximised [7, 38].

4.5.2 Community Aggregation in Louvain Method

The second step is to build new nodes that are the communities established by the previous step. Links that connect these newly generated nodes are the sum of the previously connected nodes in their corresponding communities. After constructing a new network based on these newly generated nodes, the first phase is then recalled once again to generate communities with high modularity scores [18]. Just like the first step, the second part is repeated until there is no improvement in the modularity score.

4.5.3 Community Evolution

To track down the evolution of co-investment communities over time, we used the stabilised Louvain method as proposed by Gao *et al.* [17] for detecting community changes over time in temporal evolving networks. As discussed previously, the Louvain method yields communities based on modularity gained through iterations. We faced a significant challenge while performing community detection on cumulative co-investment networks. Since generated communities change based on each iteration of our annual networks, performing community detection on each year's network, yielded different results for each year. In other words, it was impossible to keep track of communities as they were at the start of our study's time frame. A way to address this problem when detecting dynamic communities is to modify the Louvain method so that instead of running it separately on each annual network, communities detected in the first network are used as the reference point for the next year. Basically, this modification" is to change the initial partition of the network at time t to the detected partition at time *t-1*; thus, the initial partition is constrained to take into account the communities found at the previous time steps, making it possible to identify the real trends" [17]. By using this modified version of the Louvain method, it is possible to keep track of each communities' members and study the growth, shrinking or creation of new communities.

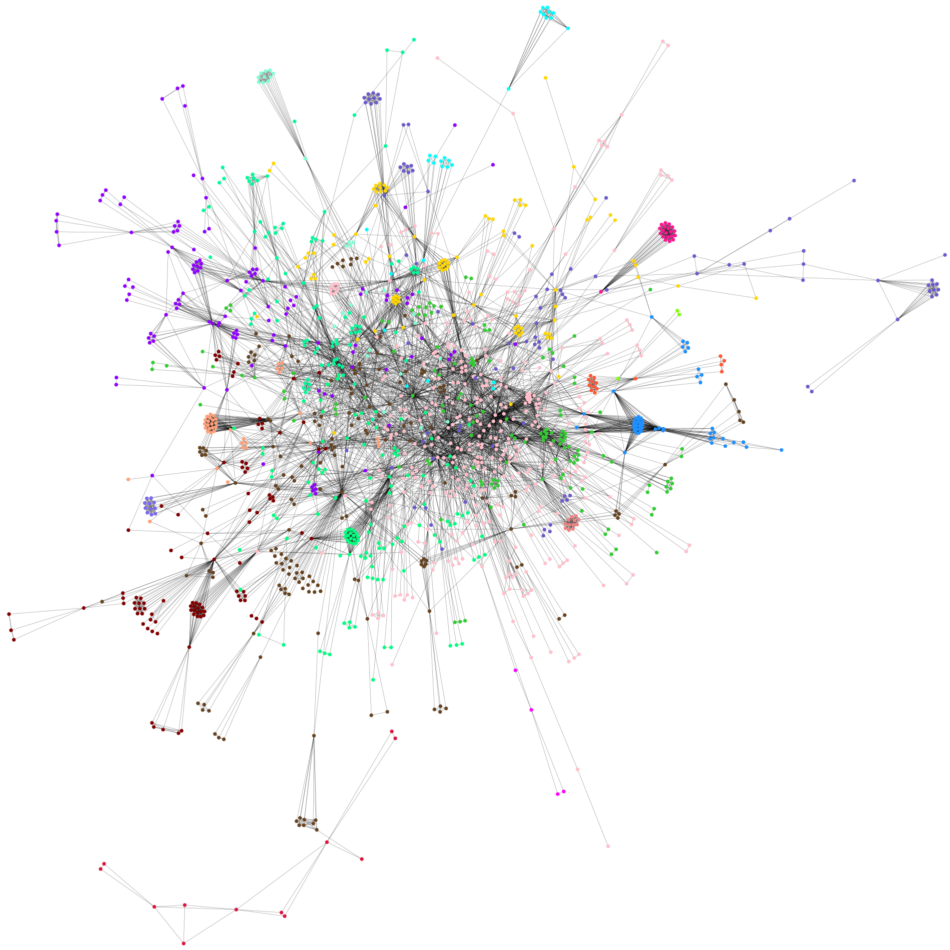

Figure 4.3: Communities indicated by colour in the GC of the co-investment network, 2021

Chapter 5

Experiments

In this chapter, we discuss the experiments conducted to answer the research question outlined in Chapter 1. To achieve this, we first conduct an exploratory analysis of our data to identify overall patterns and trends. Then we present the network metrics of the ecosystem and the co-investment networks. Moreover, we identify influential investors and look at the distribution of centrality measures across the co-investment network. We also discuss the absence or presence of preferential attachment in the growth of the co-investment network. Lastly, we identify co-investment communities between 2015 and 2021, analyse the composition of a community and track its development over the period of 2018–2021. All measures mentioned and utilised in the section were implemented using networkX [18], a Python library for studying graphs and networks.

5.1 Exploratory Analysis

This section will outline the exploratory data analysis to identify patterns, trends, and relationships that later shape our understanding of the ecosystem and co-investment networks. Moreover, this section introduces and describes a few commonly used terms and phrases in entrepreneurial finance. It is important to keep the limitations and scope of error of the dataset used in this research in mind. For more information, please look at Chapter 3.

5.1.1 Startups and Companies

Here we focus on several key aspects of the startups and companies operating in the Dutch ecosystem. All companies considered for this section had at least one round of investment between 2000–2021. First, we start by looking at Figure 5.1, which indicates the number of startups established in the period of 2000–2021 and their current operating status. As shown, despite fluctuations between 2000 and 2007, the overall trend is upward until 2015. The number of companies reached an all-time high

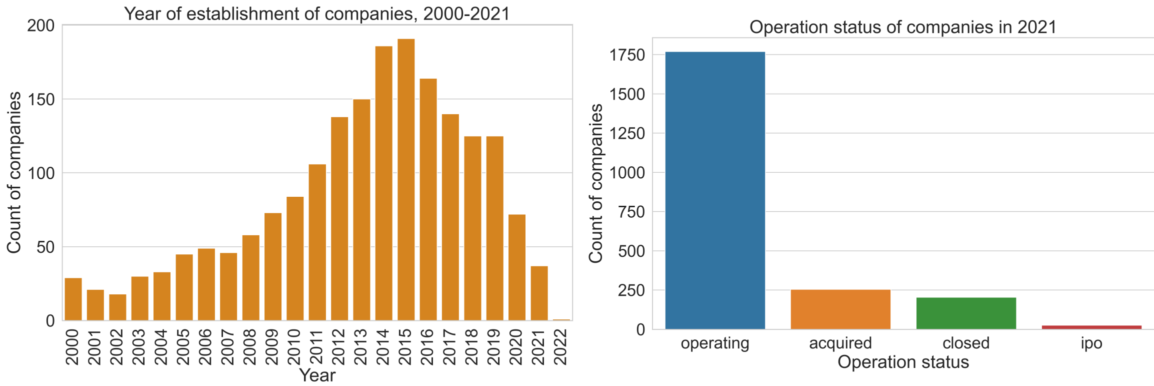

Figure 5.1: Year of establishment and operating status

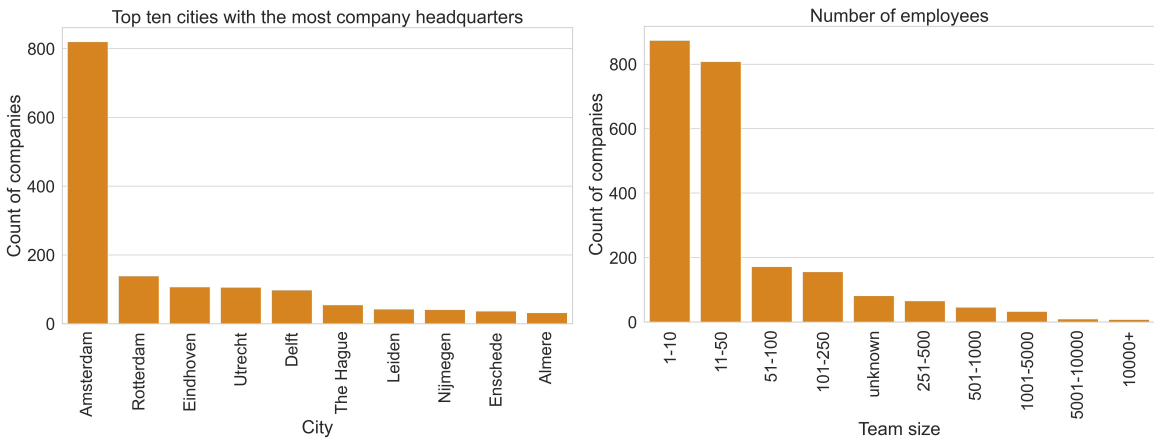

Figure 5.2: Size of teams and HQ cities

in 2015, counting over 174 that year. However, the number of startups continued to decline until 2019 and, went down sharply in 2020 and 2021. Only one startup had 2022 as its establishment year, and there are a number of startups established before 2000, but since they received investment in 2021, we kept them in the data. Almost 1,770 of the companies in our dataset are still active. As for the rest, almost 250 were acquired by other companies or startups, and 204 ceased operations. Interestingly, only 26 of the startups had an IPO. Being acquired or going public are the occasions when investors can make a return on their investment. These two processes are indicators of the success of both startups and investors in creating and selling a company. Figure 5.2 indicates the number of employees and the top ten cities where their headquarters are based. The vast majority of Dutch companies operate in Amsterdam, followed by a smaller number of companies in Eindhoven, Rotterdam and Utrecht. As for the number of employees, most of the startups have teams consisting

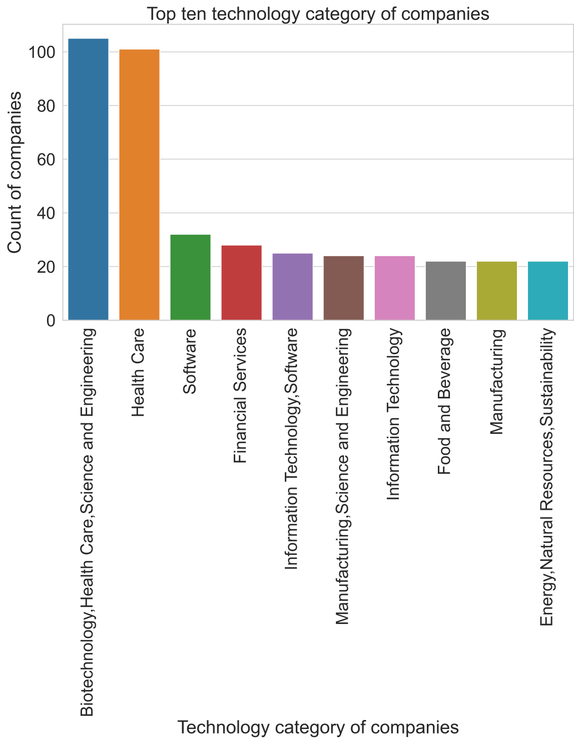

Figure 5.3: Count of top ten technology category of companies

of small-size (1–10) and medium teams (11–50). Lastly, the startup's related industries are shown in Figure 5.3. A significant majority of Dutch companies are active in biotechnology, health care, and science and engineering. Software, financial services, and IT are also common industries that Dutch startups are focused on.

5.1.2 Investors and Investments

Since this study focuses on the network of co-investments, for answering our research questions, it is important to have a general understanding of investment patterns. We start by looking at the nationality of investors involved in the Dutch ecosystem. Figure 5.4 shows the number of investments made by Dutch and international investors. We observe the prominence of Dutch investors compared to international investors. Despite the growth of British, Belgian and German investments after the 2010s, founding rounds involving American investors are the second largest group. The number of investments from both Dutch and American investors has been increasing notably since 2020. In Figure 5.5, the number of investments per year in each category of funding rounds is shown. The seed stage, the first funding round a young company receives, had the highest number of investments and grew sharply after 2019. Funding rounds with unknown series ranked second.

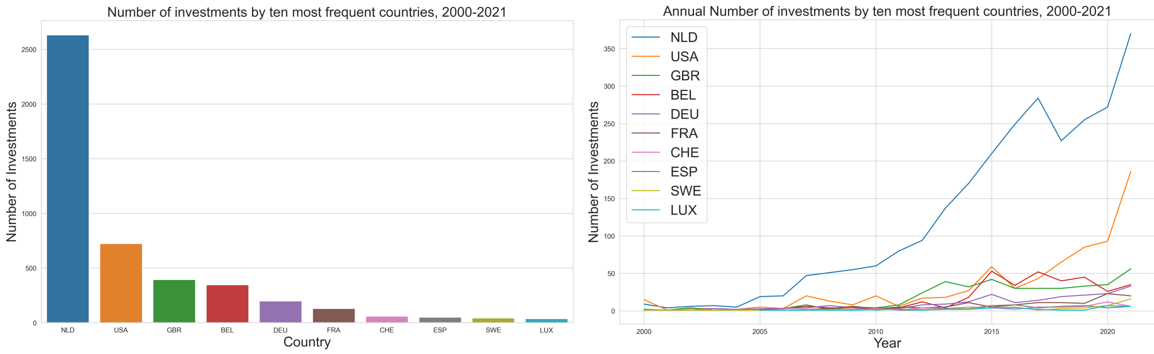

Figure 5.4: Dutch and international investments, 2000–2021

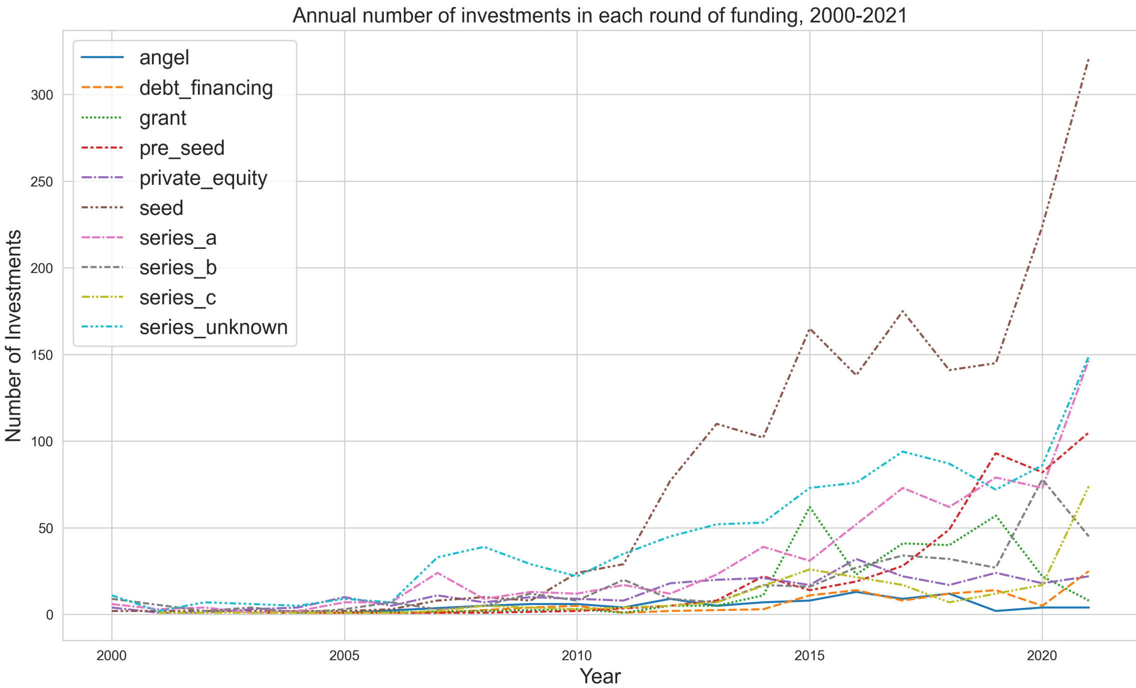

Figure 5.5: Number of investments in ten most frequent funding rounds, 2000–2021

In third place, we can see series A investments, which grew steadily over the years and saw a significant increase in 2020-2021. The number of pre-seed rounds and private equity investments also grew after 2020. The overall pattern we deduce from this figure is that early-stage funding rounds had the highest number of investments, besides private equity, and that number has continued to rise. Another important property of investors is their type. Figure 5.6 depicts the annual number of investments by the top 5 investor types. The most frequent investors are venture capital firms,

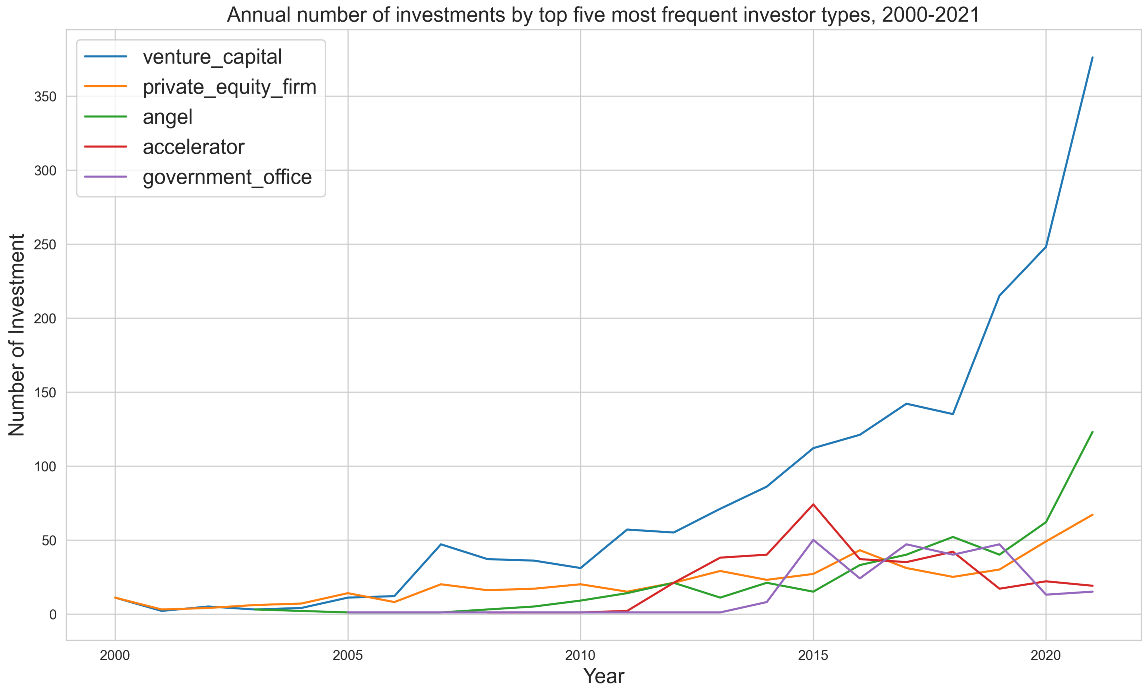

Figure 5.6: Number of investments by investor type, 2000–2021

as they are considered the main financiers of startups. We see that not only were VC firms always the most active type of investors, but their number of investments increased by three folds after 2018. The second and third most active investor types were private equity firms and angel investors, respectively. Private equity firms usually invest in more established startups and firms when less risk is associated with an investment. On the contrary, angels invest in small early rounds. The last category of investors is government offices, which are initiatives or organizations funded by government agencies. Although their share of investments is not as large as other big investors, as we will see, they play an essential role in the co-investment network. It is notable that although the number of investments grew considerably after 2019, as indicated in Figure 5.5 and 5.6, we did not witness a rising trend in the number of new startups in the same period.

5.2 Results

In this section, we start by looking at network metrics discussed in Chapter 4 for both ecosystem and co-investment networks. We then look at centrality measures and identify influential investors and discuss the distribution of centrality measures across the co-investment network. Moreover, we look at the growth of the co-investment network and discuss the presence or absence of preferential attachment in the co-investment network. Lastly, we deploy the dynamic Louvain method to identify communities of co-investment. These results will be discussed in the light of our research questions outlined in Chapter 1.

5.2.1 Network Measures

Here, we discuss the results of network metrics to answer our first research concerning the topological features of the Dutch ecosystem and co-investment network, and their evolution over time.

Network Metrics of The Two-Mode Ecosystem Network

Table 5.1 shows network metrics for the cumulative ecosystem networks for the period of 2000–2021.

Year	Nodes	Investor Nodes	Startup Nodes	Links	%Nodes in GC	%Links in GC	Density
2000	74	38	36	60	16.2	18.3	0.0219
2001	84	67	42	42	14.2	16.4	0.0199
2002	101	82	51	50	11.8	13.4	0.0161
2003	120	60	60	99	10.8	12.1	0.0138
2004	140	70	70	112	9.2	10.7	0.0114
2005	182	90	92	145	7.6	8.9	0.0088
2006	212	105	107	177	11.3	14.6	0.0079
2007	314	160	154	272	15.6	20.2	0.0055
2008	396	191	205	359	20.4	27	0.0046
2009	484	229	255	452	32.2	43.5	0.0039
2010	579	274	305	554	37.4	49.2	0.0033
2011	686	322	364	674	39.3	51.3	0.0029
2012	874	410	464	865	36.7	46.9	0.0023
2013	1,104	502	602	1,112	36.2	45.2	0.0018
2014	1,363	608	755	1,416	49.9	62.2	0.0015
2015	1,720	740	980	1,863	56.6	68.7	0.0013
2016	2,089	893	1,196	2,301	62.8	75.3	0.0011
2017	2,529	1,066	1,463	2,832	66.2	78.4	0.0009
2018	2,927	1,264	1,663	3,336	66.4	78.6	0.0008
2019	3,353	1,470	1,883	3,914	68.8	80.6	0.0007
2020	3,769	1,708	2,061	4,566	72.6	83.8	0.0006
2021	4,426	2,156	2,270	5,540	74.3	85	0.0006

Table 5.1: Network metrics of the ecosystem networks, 2000–2021

As we can see in Table 5.1, the ecosystem network grew significantly both in terms of nodes and links between 2000–2021. Although the ecosystem's growth is notable, it is important to keep in mind that this increase in the number of nodes and links is cumulative. In the year 2000, only 16% of nodes and 18% of links were a part of the GC. However, in the year 2021, these figures rose to 74% and 85%, respectively. This trend indicates that the network became more connected over this period. Moreover, until 2007, the network consisted of roughly an equal number of investors and funded startups. However, after this year, we see more startups compared to investors in the

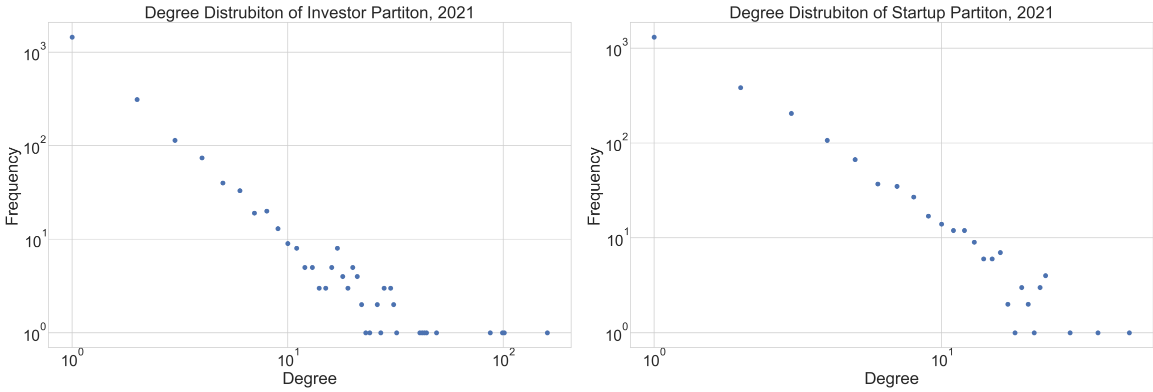

Figure 5.7: Degree distribution of the investor partition (left) and the startup partition (right)

network. Lastly, we look at the network's density. Since the beginning of the time window of this study, networks have been sparse. In the year 2000, the network's density was 0.02, indicating that only two per cent of possible investments were made in the network. As we move forward to 2021, the network becomes extremely sparse, with a density value of close to zero.

The degree distributions of the investor partition and startup partition are shown in Figure 5.7, depicting how often a degree score occurs in the ecosystem network. Both distributions have a heavy-tailed distribution, meaning that most investors had a small number of investments. On the startup side, a large portion of them did not receive many investments. However, in both the startup and investor partition, a small number of investors and startups were considerably more active than the rest of their peers.

Network Metrics of the One-Mode Co-investment Network

Now that we looked briefly at the ecosystem network metrics, we move to a general overview of co-investment networks. We start by examining the number of nodes and links, and the percentage of them being a part of the GC. The network's number of investors and co-investments rose sharply between 2000 and 2021. In 2021, 70% of investors and 90% of co-investments were present in the GC, compared to only 18% and 40% in 2000. By looking at the average weighted degree of investors, we can see a gradual increase in the average number of co-investments from less than three during the 2000s to just below eight in 2021, again indicating the incremental movement of the network to a more connected state. Furthermore, we can see an increase in the average shortest path from 1.3 to 4.06 over the time period of this study. This measure indicates that for investors to be connected through co-investments, 4 links are expected to be traversed on average. To get a better sense of the size of the network, we looked at the diameter. The diameter of the network grew considerably from two in 2000 to thirteen in 2021. The overall upward trend of diameter and average shortest path length indicates the growing distance between investors and the sparsity of the network.

Year	Nodes	Links	% Nodes in GC	% Links in GC	Diameter GC	ASPL	Density	Average Clustering Coefficient	Transitivity	Average Degree	Degree Assortativity Coefficient
2000	38	32	18.4	40.6	2	1.38	0.045	0.470	0.78	1.6	0.405
2001	42	33	16.6	30.3	2	1.52	0.038	0.422	0.73	1.5	0.212
2002	51	43	13.7	23.2	2	1.52	0.034	0.472	0.75	1.6	0.226
2003	60	50	11.6	20	2	1.52	0.028	0.468	0.78	1.6	0.262
2004	70	54	10	18.5	2	1.52	0.022	0.444	0.79	1.5	0.313
2005	90	72	11.1	22.2	3	1.91	0.018	0.390	0.74	1.6	0.279
2006	105	96	16.1	29.1	4	2.33	0.018	0.453	0.69	1.8	0.205
2007	160	188	19.3	28.7	6	3.12	0.015	0.241	0.78	2.3	0.579
2008	191	248	23.5	39.5	8	3.70	0.014	0.160	0.78	2.6	0.540
2009	229	307	35.3	58.9	13	6.11	0.012	0.155	0.72	2.7	0.376
2010	274	376	38.6	61.7	14	5.96	0.010	0.157	0.66	2.8	0.230
2011	322	500	39.7	65.6	11	4.68	0.010	0.164	0.66	3.1	0.346
2012	410	697	36.5	57.1	11	4.82	0.008	0.126	0.69	3.4	0.330
2013	502	873	33.4	49.9	12	5.05	0.007	0.042	0.67	3.6	0.176
2014	608	1,138	45	73.3	12	5.26	0.006	0.031	0.63	3.9	0.212
2015	740	1,620	51.8	82.2	11	4.94	0.006	0.032	0.61	4.6	0.095
2016	893	2,155	57.1	88.1	12	4.68	0.005	0.032	0.59	5.1	0.056
2017	1,066	2,682	59.7	90	11	4.44	0.005	0.033	0.53	5.3	0.036
2018	1,264	3,403	60.3	88.8	11	4.17	0.004	0.034	0.51	5.7	0.048
2019	1,470	4,149	62.5	89.2	10	4.03	0.004	0.036	0.46	5.9	0.018
2020	1,710	5,436	67.8	88.4	13	4.09	0.004	0.037	0.44	6.7	0.010
2021	2,159	7,996	70.9	90.6	13	4.06	0.003	0.040	0.45	7.7	0.021

Table 5.2: Network metrics of the co-investment networks, 2000–2021

The average clustering of the network saw a significant decrease as more co-investments were made. The average clustering coefficient for the 2021 cumulative network is 0.04, indicating the low probability of an existing co-investment between two neighbours of a randomly selected investor in the network. Another clustering measure, transitivity, also indicates the same trend. The transitivity of the network fell from 0.7 in 2000 to 0.4 in 2021, pointing to the fact that only 40% of the possible closed triangles are presented in the network. Moreover, the decline in the network's density also suggests that the network has become less dense and clustered over the years. Additionally, the degree assortativity coefficient score for the network in 2021, is 0.021. The declining score of the degree assortativity coefficient indicates that investors do not necessarily make new co-investments based on the number of shared investments of their candidate partners. In other words, we established that assortativity by degree did not play a role in shaping co-investments in 2021. However, in the early years of this study's time window, between 2000–2008, the associativity was considerably higher, indicating the existence of the tendency of investors with a high number of co-investments to share funding rounds with less active peers. We look more closely at the tendency of investors to form a connection base on degree later in this Chapter.

Lastly, we looked at the occurrence of different degree scores in the network. Figure 5.8 shows the network's degree distribution in 2021, indicating a heavy-tailed distribution. In the case of this network, this signifies that there are a few investors with a high number of co-investments, but that most nodes have a considerably lower number of co-investments, most frequently one or two.

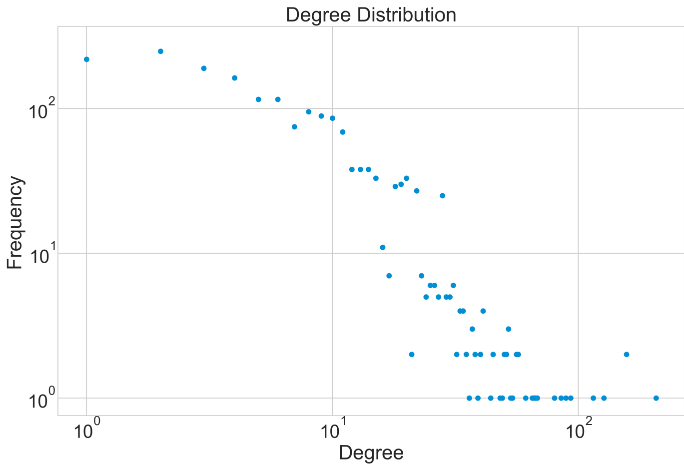

Figure 5.8: Degree distribution co-investment network, 2021

5.2.2 Centrality Measures

In this section we explore the centrality measured previously discussed in Chapter 4. First, we identify investors with the highest centrality scores and then present the distribution of centrality measures for the 2021 co-investment network.

Degree Centrality

	Investor	Centrality Score	Type	Country
1	EASME - EU Executive Agency for SMEs	0.07	Government Office	BEL
2	BOM	0.05	Government Office,Venture Capital	NLD
3	Rockstart	0.05	Accelerator,Venture Capital	NLD
4	Inkef	0.05	Venture Capital	NLD
5	Startupbootcamp	0.04	Accelerator	GBR

Table 5.3: Investors with highest degree centrality scores, 2021

By using the degree centrality, we can examine the number of co-investments by an investor. We start by identifying five investors with the highest degree centrality score in the network, as shown in Table 5.3. In this network, high degree centrality scores for the top five investors demonstrate the high number of co-investment these investors made in the Dutch ecosystem. The Executive Agency for Small and Medium-sized Enterprises (EASME), an EU agency based in Belgium, ranked highest with a degree centrality score of 0.07, meaning that the EASME was involved with 7% of all co-investments made. Brabantse Ontwikkelings Maatschappij (BOM), another government agency based in the Province of Brabant, Rockstart, an accelerator, and Inkef, a Dutch VC, all scored 0.05. Lastly is the startupbootcamp, a British accelerator that was involved in co-investments with 4% of other investors. As we look at these figures, a pattern starts to emerge. All top five investors are government offices and/or early-stage investors. For instance, EASME's most frequent type of investment was 152 grants to startups. Moreover, BOM's primary investments were in seed, Series A and B, and, Rockstart invested 86 times in seed rounds and twelve times in pre-seed out of its total of 102 investments. We argue that the high centrality score of these investors is due to the fact that when they made a high number of early investments in startups, those startups then went on to attract more investments, resulting in a higher number of their co-investments. Figure 5.9 depicts the

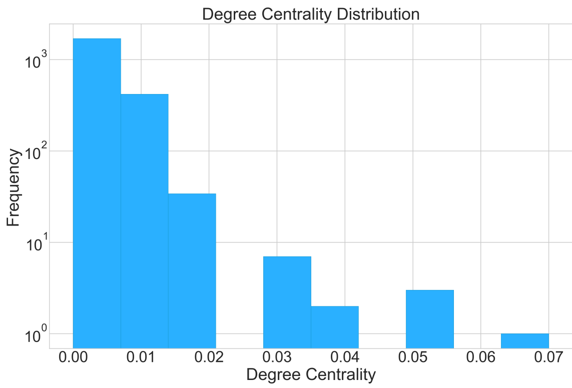

Figure 5.9: Distribution of degree centrality scores, 2021

degree centrality histogram for the 2021 cumulative network. Most investors have a degree centrality of less than 0.02. The number of investors with a relatively higher degree centrality than the rest is very small. This observation suggests that this network gives more weight to nodes that share co-investments with other investors. Moreover, investors who either did not share their investments with others or were relatively inactive scored low degree centrality scores.

Betweenness Centrality

	Investor	Centrality Score	Type	Country
1	Rockstart	0.09	Accelerator,Venture Capital	NLD
2	EASME - EU Executive Agency for SMEs	0.08	Government Office	BEL
3	Startupbootcamp	0.05	Accelerator	GBR
4	Inkef	0.05	Venture Capital	NLD
5	BOM	0.04	Government Office, Venture Capital	NLD

Table 5.4: Investors with highest betweenness centrality scores, 2021

An investor with a high betweenness centrality score acts as a network bridge or broker. By studying Table 5.4, we see the same investors with the highest degree centrality score, although the ordering is different. To explain what betweenness centrality indicates, we examine investors that scored amount the top 5 based on this measure. For example, Rockstart, which scored highest, acted as an important bridge between disconnected investors. The same goes for the rest of the investors with high betweenness scores, indicating that these investors participated in funding rounds that involved coordinating and the brokerage of other investment rounds. Since these five investors scored the highest on both degree and betweenness measures, they shared the most co-investments and were in an influential position to connect disconnected investors and also control information flow between other investors in the network. Figure 5.4 indicates most investors yield very low scores of

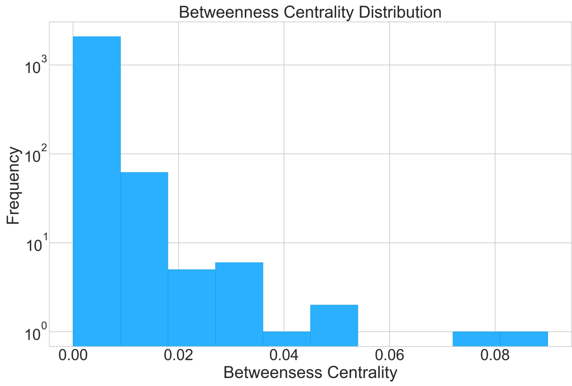

Figure 5.10: Distribution of betweenness centrality scores, 2021

betweenness centrality. 2,159 nodes are present in this network, of which 2,081 nodes scored zero on betweenness centrality, and only 62 scored 0.01. This result indicates that this network is extremely sparse, connected by a small number of investors with a high number of co-investments.

Eigenvector Centrality

	Investor	Centrality Score	Type	Country
1	Startupbootcamp	0.44	Accelerator	GBR
2	EASME - EU Executive Agency for SMEs	0.33	Government Office	BEL
3	BOM	0.33	Government Office, Venture Capital	NLD
4	Startupbootcamp HighTechXL	0.24	Accelerator	NLD
5	Startupbootcamp Amsterdam	0.23	Accelerator, Private Equity Firm, Venture Capital	NLD

Table 5.5: Investors with highest eigenvector centrality scores, 2021

As discussed before, eigenvector centrality measures an investor's centrality as a proportion of the number of its neighbour's co-investments. Interestingly, three of the top five investors with the highest eigenvector centrality scores are bootcamps. Bootcamps are among the earliest organisations helping startups to start working on their ideas and businesses. Besides Startupbootcamp and EASME, which scored high on degree centrality, Startupbootcamp HighTechXL and Startupbootcamp Amsterdam scored 0.02 and 0.01 on degree centrality scores, respectively. This finding indicates that these two investors did not share many investors, but those investments were made with other highly connected investors. Startupbootcamp scored highest with an eigenvector centrality score of 0.44; if we consider that it also had some of the highest degree and betweenness scores, and two other bootcamps in this list are its subsidiaries, it is visible that Startupbootcamp is one of the most influential nodes in the co-investment network. We should mention that Startupbootcamp HighTechXL (in Eindhoven), and Startupbootcamp Amsterdam are all subsidiaries or programs of the same entity, namely Startupbootcamp (based in London). However, in our dataset, they were indicated as different entities. Hence, we treat them as separate entities. As we can see in Figure 5.11, most of the investors scored zero on this measure. A significant number of investors fall into a range of 0.01 to 0.05. However, a small number of investors with relatively high eigenvector centrality scores are present in the network.

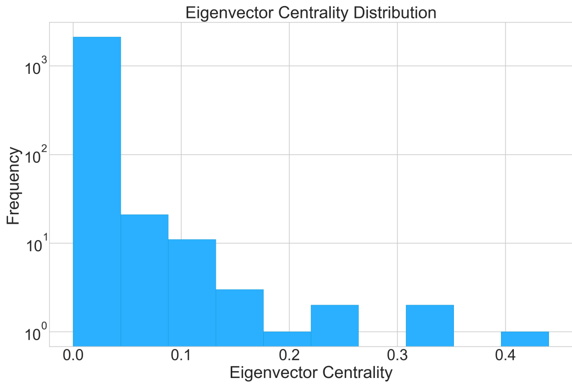

Figure 5.11: Distribution of eigenvector centrality scores, 2021

5.2.3 Preferential Attachment

In this section, we aim to answer one of the research questions of this research, namely explaining the mechanisms behind the growth of the co-investment network. To answer this question, we study the presence or absence of preferential attachment (PA). We start by analysing Figure 5.12. Across all four years, investors with a small number of co-investments gained the highest number of new co-investments. However, the overall pattern is that investors with a substantial number of co-investments did not gain many new co-investments. These findings suggest the first evidence against the existence of PA, since we do not observe that highly active investors gain most of the new connections. As we can see, Figure 5.12 is noisy: therefore, it is difficult to establish the exact type of network growth using the mentioned results. Figure 5.13 shows the double-logarithmic plot of cumulative growth in degree for four consecutive years, 2017–2020. On the horizontal axis, the weighted degrees for each corresponding year are plotted; on the vertical axis, we see the cumulative sum of degree gain for nodes with a given degree. Although this method is not an exact replication of Barabási's proposed method [4], studying the plots in Figure 5.13 makes it possible to analyse PA in co-investment networks. We provided three lines to guide the eye, indicating the existence of PA, linear growth and reverse PA for the growth of the network. If the cumulative growth of the degree corresponds to the linear PA line, we can detect the PA. Otherwise, if this line falls below the linear PA line, we argue for the absence of PA. In all four annual periods, it is visible that the cumulative growth of investors with a high number of shared investments is sublinear and between random and reverse PA lines. For example, in the period of 2020–2021, investors with a relatively low number of shared investments gained cumulative degree growth in accordance with the random model. However, this degree of growth for high-degree investors becomes did not continue at this rate and even fell under the rate of the random regime. This pattern is visible across all four snapshots of the network we studied. Therefore, we found empirical evidence against the existence of PA to explain the growth of the co-investment network.

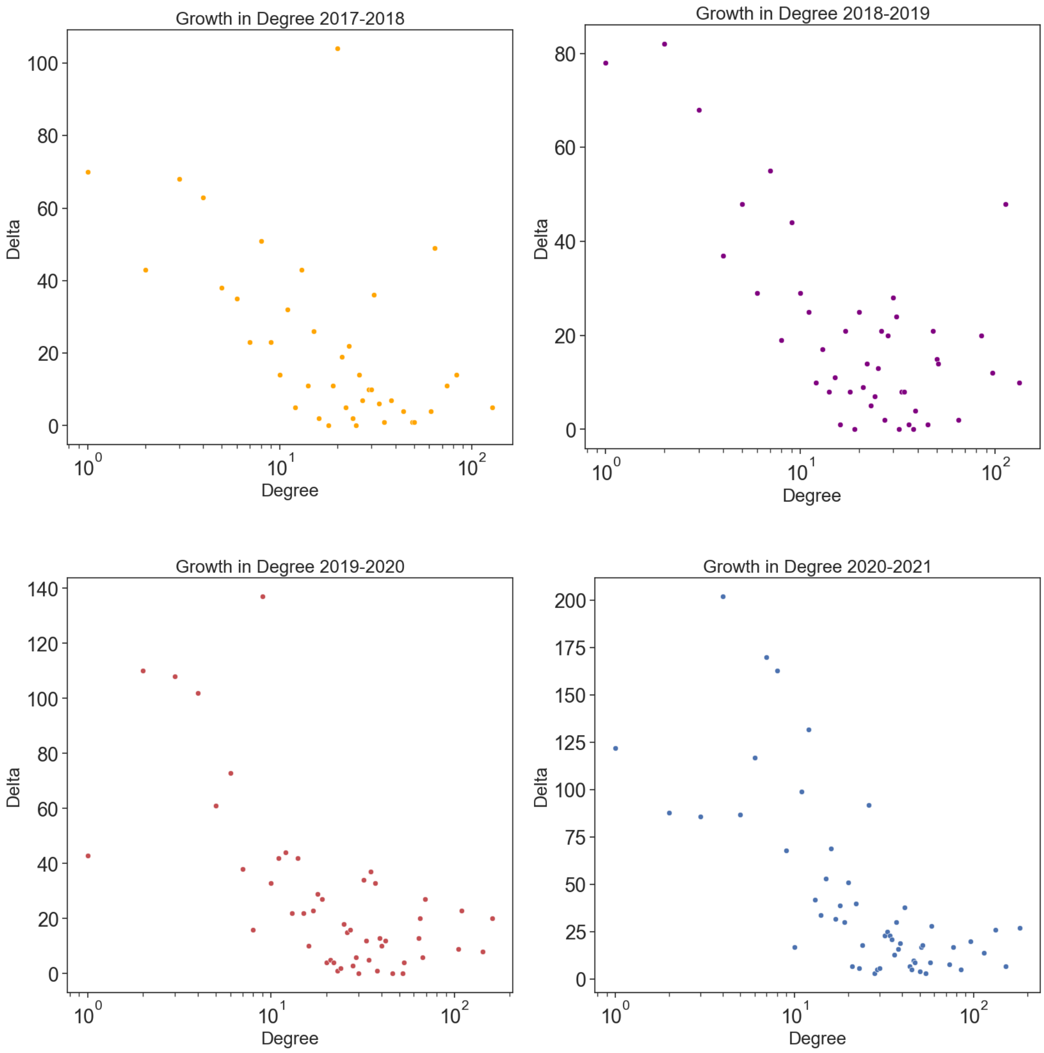

Figure 5.12: Gain in degree between two consecutive years, 2017-2020

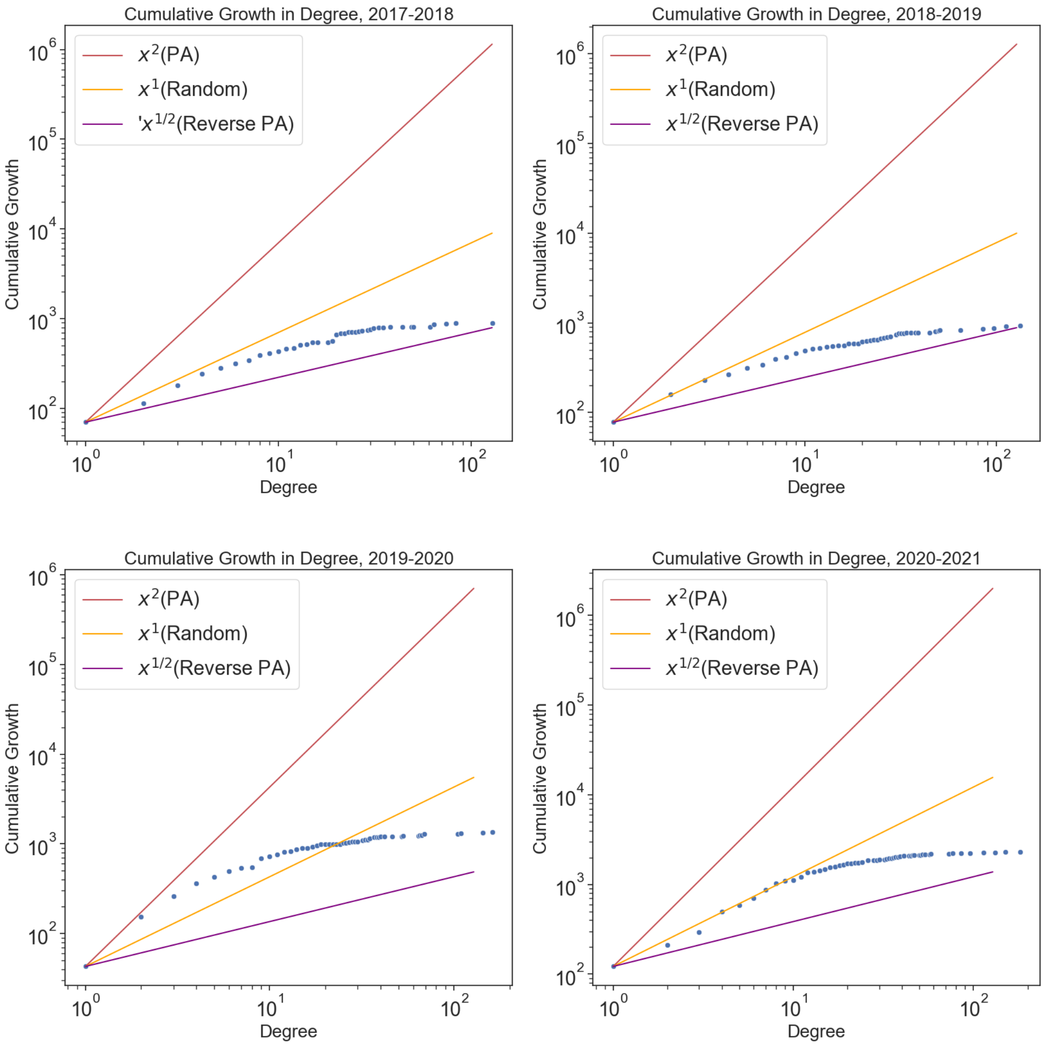

Figure 5.13: Cumulative growth in degree between two consecutive years, 2017-2020

5.2.4 Community Detection

In this section, we identify investors' communities using the modified Louvain method discussed in Chapter 4. Although it was possible to start community detection from the year 2000, we chose only to study communities between 2015–2021. We decided to set the starting point of community detection in 2015 to be able to track communities present in 2021. Lastly, we ran the dynamic Louvain method only on the GC of each year. As we can see in Table 5.6, the number of communities detected in this period first fluctuated between fourteen and fifteen communities from 2015 to 2018. Then, one community was added yearly until, in 2021, nineteen communities were present.

Year	Communities
2015	14
2016	15
2017	14
2018	15
2019	17
2020	18
2021	19

Table 5.6: Number of communities in the GC, 2015–2021

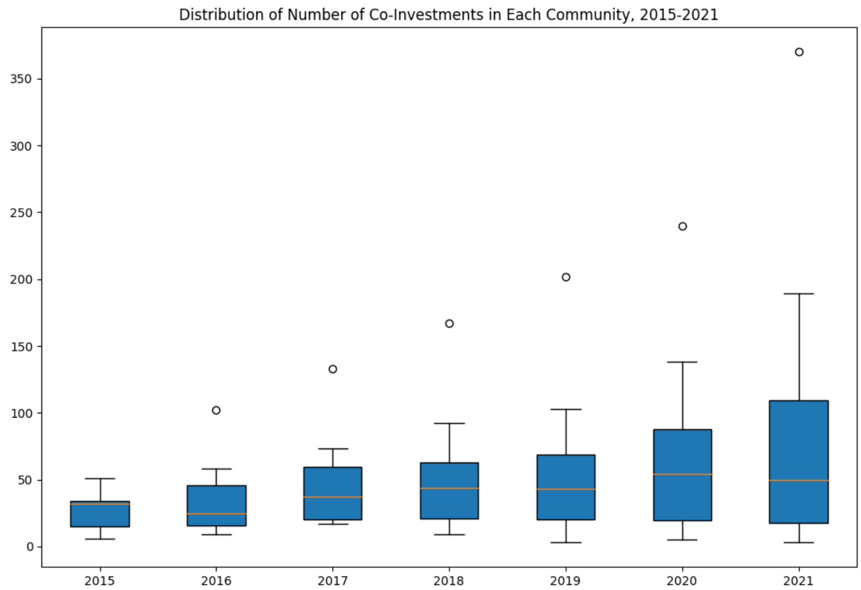

Figure 5.14: Distribution of the number of co-investments in communities of each year, 2015–2021

To better understand the general characteristics of these communities, it is useful to look at the number of co-investments made inside of them. As indicated in Figure 5.14, there are communities with a small number of members each year. From 2016 onward, each year, an outlier community with a significantly higher number of co-investments than the rest was present. In 2016, this outlier had 102 members, and in 2021, it grew to 370 investors. Besides the smallest and biggest communities, it is visible in Figure 5.14 that the lower quartile of the number of members of communities across the years remained roughly the same. Still, the upper quartile and upper extreme grew steadily across this time window. Spread within the second and third quartiles increased over the years, and although the median fluctuated, it moved from 32 co-investments in 2015 to 50 in 2021.

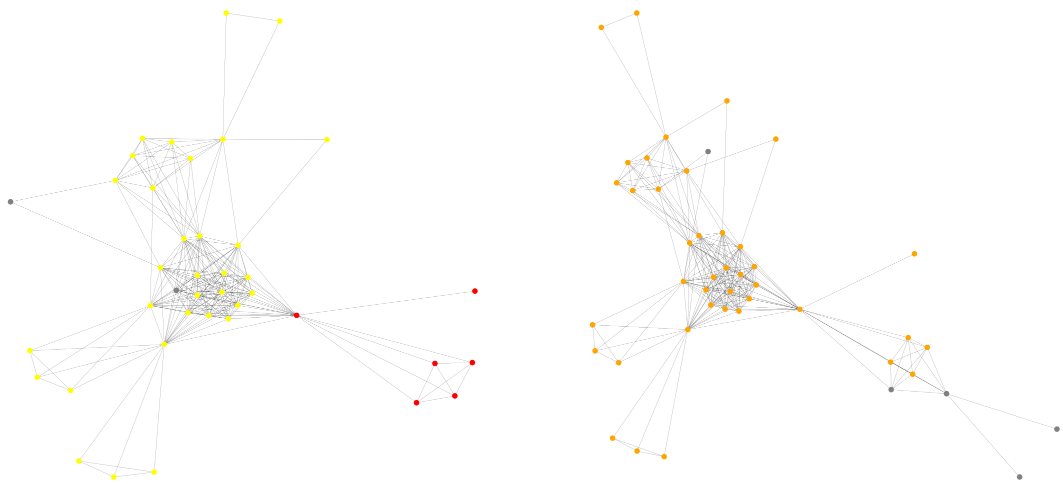

Figure 5.15: Left: our community of interest in 2016. Yellow and red nodes existed as two separate communities in 2015. Right: the community division in 2017. Orange nodes were already present in 2016, and grey nodes were added in 2017

Furthermore, we look at the composition of a community and track its development over the years. We start by considering the smallest community of 2015, with only six members, and track down its evolution. This community had six members, indicated as red nodes in Figure 5.15 in 2015. This community consisted of Belgian, Dutch, and one German investor. In terms of the type of investors, five are VCs, and one is a private equity firm. Moreover, we tracked down their prior investments, and it was indicated that they all shared a series B prior co-investments in an Eindhoven-based startup in 2015. Although sharing this investment formed this community, two of them, a German and a Dutch VC, had already previously shared a series A investment in the same startup. This small community was combined with a bigger community of 32 members in 2015, represented as yellow nodes in Figure 5.15, to make a bigger community of 46 investors in 2016. This community steadily grew to 73 members in 2020, as observed in Figures 5.16 and 5.17. Interestingly, in 2021 by combining with another big community, this group of investors reached 370 and ranked as the biggest community in 2021. By looking at this example, it is possible to verify that the identified community is built upon the prior ties of its members. Moreover, throughout the yearly evolution of the ecosystem, new communities are born, whether via new co-investments by recently added investors to the network or the combination of prior communities based on the co-investments of their members. Overall, besides the prior ties that investors shape in their syndication, there is no clear pattern to explain if communities are based on nationality or type of investors or even if the round of investment is a determining factor.

Figure 5.16: Left: the community division in 2018. Right: the community division in 2019. Orange nodes were already present in the year before and grey nodes were added in the indicated year

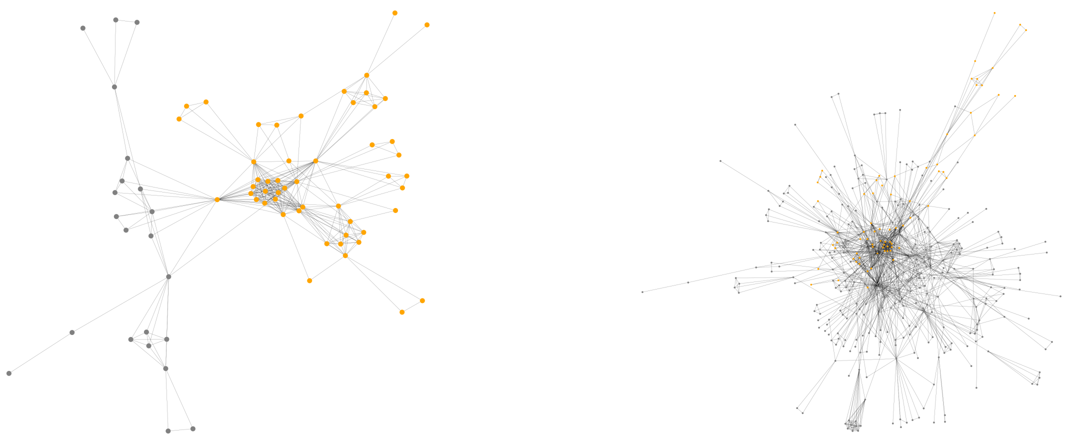

Figure 5.17: Left: the community division in 2020. Right: the community division in 2021. Orange nodes were already present in the year before, and grey nodes were added in the indicated year

The structure of communities found in 2021 follows the general trends of investments that we mentioned before. For example, in the three biggest communities of 2021, the most frequent funding rounds were pre-seed, seed, series A and B and unknown series. Moreover, as we discussed in Chapter 2, investing in a similar industry is an important motivation for foreign investors trying to select other domestic or international investors to invest with, we also analysed the technology category of investments. As for our biggest community of 2021, with 1,704 rounds of investment, the most frequent technology categories were biotechnology and healthcare. However, there is a wide range of other industries that investors from this community invested in, ranging from manufacturing, science and engineering to energy and sustainability. We saw this investment pattern again in the third biggest community with 139 investors and 434 rounds of investment, indicating the general prevalence of biotechnology and health care. However, the second biggest community, with 180 investors and 480 investments, did not indicate any meaningful pattern of investment in a particular industry. Overall, we did not find any indication pointing out the fact that communities of investors are built around an interest in any technology sector.

Chapter 6

Conclusion and Future Work

In this section, we provide our findings regarding the research questions of this thesis.
We explored several topological aspects of the ecosystem and co-investment networks and their evolution throughout the years. We concluded that although the co-investment network grew substantially in size and connectedness, it became considerably less clustered and dense and more sparse. As for our second research question, our findings indicated the prominence of early–stage investors in the Dutch co-investment network. Dutch and international government offices, accelerators, bootcamps and early-stage venture capital firms were the most important players in the co-investment network. If we keep in mind that these investors systematically invest in early funding rounds, we deduce why they ranked high on all centrality measures. As for explaining the growth mechanism of the co-investment network, we found the absence of preferential attachment in the co-investment network. Lastly, we looked at the process of community formation and evolution. We found that communities of co-investment are built around shared investments in one or more startups, and their number of members changed significantly over time, either by gaining new members or merging with other communities. We did not find any evidence of other determining factors on community formation, such as the investors' nationality, type or the sector of their investments.

By performing an exploratory analysis of our data, we provided context for the understanding the implications of our network analysis. We showed that the number of new startups with funding in the Dutch EE grew substantially in the period of 2000–2015. Although this number went down until the end of 2021, we need to keep in mind the effect of the Covid-19 pandemic and also the fact that we only considered startups with at least one round of funding. Moreover, we showed that while Amsterdam is the biggest hub of entrepreneurial activity, other important cities have a substantial number of startups. As for investments, we encountered an ever-growing trend of international investors, most notably American investor involvement, especially after 2020. We identified early-stage investments as the most frequently invested rounds. Although late-stage funding rounds were also present in the top ten rounds of investments, the lack of investment in late-stage rounds compared to early–stage investments is an important finding. This discovery is in line with a recent report by Techleap [37] that indicated the startup-to-scaleup ratio in the Netherlands is behind its competitors in Germany, the UK and France. Based on these results, from a policy point of view, we believe that more efforts should be made to encourage international investors in late-stage funding. The presence of specialised international and domestic investors will help the Dutch EE to fully realise its potential.

We believe that our research created a comprehensive network-based approach to the study of co-investment networks. Our research can be extended in a few ways. First, we did not include the amount of money raised in each round of funding. By including this information, it will be possible to include weights for links in the ecosystem network and improve the accuracy of several measures deployed on the networks in question. Secondly, for constructing the cumulative ecosystem networks, the inclusion of a time-series analysis of IPOs, acquisitions and closure of startups will open up a different way of approaching a co-investments network. These networks can include investments that only occurred in a given year or month. On a different note, new research can be dedicated to evaluating the performance of various community detection methods while empirically analysing the structure of communities. Moreover, as we indicated the absence of evidence on the existence of preferential attachment, future research will be beneficial to explain why active investors do not continue to co-invest at a substantial rate. More importantly, future work can also focus on the contagion and diffusion of financial shocks in the financial networks supporting startups in an EE. At the time of writing this thesis, the takeover of Silicon Valley Bank shook the world of startups and entrepreneurship finance. After the financial crisis of 2008, network models for studying the spread and diffusion of failures of interconnected banks were developed and deployed. The same approach was again used to study the effects of supply chain failures during the Covid-19 pandemic. Considering the importance of the startup economy to developed and emerging economies, it is of vital importance to understand and study the less studied financial networks that support entrepreneurial ecosystems.

Chapter 7

References

[1] Janna Alvedalen and Ron Boschma. “A critical review of entrepreneurial ecosystems research: towards a future research agenda”. In: *European Planning Studies* 25.6 (2017), pp. 887–903.

[2] W. Brian Arthur. “Foundations of complexity economics”. In: *Nature Reviews Physics* 3.2 (2021), pp. 136–145.

[3] Albert-László Barabási and Réka Albert. “Emergence of Scaling in Random Networks”. In: *Science* 286.5439 (1999), pp. 509–512.

[4] Albert-László Barabási and Márton Pósfai. “Network science”. In: Cambridge University Press, 2016.

[5] Cristiano Bellavitis, Joost Rietveld, and Igor Filatotchev. “The effects of prior co-investments on the performance of venture capitalist syndicates: A relational agency perspective”. In: *Strategic Entrepreneurship Journal* 14.2 (2020), pp. 240–264.

[6] Matthijs den Besten. “Crunchbase Research: Monitoring Entrepreneurship Research in the age of big data”. In: *SSRN Electronic Journal: 3724395* (2020).

[7] Francis Bloch and Matthew O. Jackson. “Centrality Measures in Networks”. In: *SSRN Electronic Journal: 2749124* (2016).

[8] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. “Fast unfolding of communities in large networks”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2008.10 (2008), P10008.

[9] Phillip Bonacich. “Power and Centrality: A Family of Measures”. In: *American Journal of Sociology* 92.5 (1987), pp. 1170–1182.

[10] Ulrik Brandes. “A faster algorithm for betweenness centrality”. In: *The Journal of Mathematical Sociology* 25.2 (2001), pp. 163–177.

[11] Ulrik Brandes. “On variants of shortest-path betweenness centrality and their generic computation”. In: *Social Networks* 30.2 (2008), pp. 136–145.

[12] Crunchbase. *Crunchbase*. 2023. URL: `https://www.crunchbase.com/`(visisted on Feb 2023).

[13] Jean-Michel Dalle, Matthijs den Besten, and Carlo Menon. “Using Crunchbase for economic and managerial research”. In: *”OECD Science, Technology and Industry Working Papers, No. 2017/08* (2017).

[14] Igor Filatotchev, Mike Wright, and Mufit Arberk. “Venture capitalists, syndication and governance in initial public offerings”. In: *Small Business Economics* 26.4 (2006), pp. 337–350.

[15] Jacob G. Foster, David V. Foster, Peter Grassberger, and Maya Paczuski. “Edge direction and the structure of networks”. In: *Proceedings of the National Academy of Sciences* 107.24 (2010), pp. 10815–10820.

[16] Linton C. Freeman. “A Set of Measures of Centrality Based on Betweenness”. In: *Sociometry* 40.1 (1977), pp. 35–41.

[17] Yuan Gao, Zhen Zhu, Raja Kali, and Massimo Riccaboni. “Community evolution in patent networks: technological change and network dynamics”. In: *Applied Network Science* 3.1 (2018), p. 26.

[18] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. “Exploring Network Structure, Dynamics, and Function using NetworkX”. In: *Proceedings of the 7th Python in Science Conference*. Ed. by Gaël Varoquaux, Travis Vaught, and Jarrod Millman. Pasadena, CA USA, 2008, pp. 11–15.

[19] Tobias Henz, Laura Hofstee, Peter Jacobs, Mohcine Ouass, and Sven Smit. *Building a world-class Dutch start-up ecosystem*. Technical Report. McKinsey & Company, 2022.

[20] Yael V. Hochberg, Laura A. Lindsey, and Mark M. Westerfield. “Resource accumulation through economic ties: Evidence from venture capital”. In: *Journal of Financial Economics* 118.2 (2015), pp. 245–267.

[21] Yael V. Hochberg, Alexander Ljungqvist, and Yang Lu. “Whom you know matters: Venture capital networks and investment performance”. In: *Journal of Finance* 62.1 (2007), pp. 251–301.

[22] Christian Hopp and Finn Rieder. “What drives venture capital syndication?” In: *Applied Economics* 43.23 (2011), pp. 3089–3102.

[23] Matthew O. Jackson. *Social and Economic Networks*. Princeton University Press, 2010.

[24] Éanna Kelly, Kit Gillet, Kai Nicol-schwarz, and Ross Williams. *The future's orange? How the Netherlands can become a tech master*. Technical Report. Sifted, 2021.

[25] Susanna Khavul and David Deeds. “The Evolution of Initial Co-investment Syndications in an Emerging Venture Capital Market”. In: *Journal of International Management* 22.3 (2016), pp. 280–293.

[26] Ji Youn Kim and Haemin Dennis Park. “The influence of venture capital syndicate size on venture performance”. In: *Venture Capital* 23.2 (2021), pp. 179–203.

[27] Jiaqi Liu, Xuerong Li, Linyuan Lu, Jichang Dong, and Jinhu Lu. “Evaluating Performances and Importance of Venture Capitals: A Complex Network Approach”. In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 68.5 (2021), pp. 2060–2068.

[28] Sebastian Mallaby. *The Power Law: Venture Capital and the Making of the New Future*. Penguin Press, 2022.

[29] Colin Mason and Ross Brown. "Entrepreneurial Ecosystems and Growth-Oriented Entrepreneurship". In: Background paper prepared for the workshop organised by the OECD LEED Programme and the Dutch Ministry of Economic Affairs. 2014, pp. 1–38.

[30] Ramana Nanda and Matthew Rhodes-Kropf. *Coordination Frictions in Venture Capital Syndicates*. Working Paper 24517. National Bureau of Economic Research, 2018.

[31] M. E. J. Newman. "Mixing patterns in networks". In: *Physical Review E* 67.2 (2003).

[32] M. E. J. Newman. *Networks*. Oxford University Press, 2018.

[33] M. E. J. Newman and M. Girvan. "Finding and evaluating community structure in networks". In: *Physical Review E* 69.2 (2004), p. 026113.

[34] Thomas Schank and Dorothea Wagner. "Approximating Clustering Coefficient and Transitivity". In: *Journal of Graph Algorithms and Applications* 9.2 (2005), pp. 265–275.

[35] Bas van der Starre, Elmar Cloosterman, Dolfine Kosters, and Bella Suwarso. *The Potential of Emerging Technology in the Netherlands Report*. Technical Report. Techleap, 2021.

[36] Techleap. *Dutch Healthtech 2021 Report: unlocking its untapped potential*. Technical Report. Techleap, 2021.

[37] Techleap. *State of Dutch Tech 2023*. Technical Report. Techleap, 2023.

[38] V. A. Traag, L. Waltman, and N. J. van Eck. "From Louvain to Leiden: guaranteeing well-connected communities". In: *Scientific Reports* 9.1 (2019), p. 5233.

[39] Buvaneshwaran Venugopal and Vijay Yerramilli. "Seed-Stage Success and Growth of Angel Co-investment Networks". In: *The Review of Corporate Finance Studies* 11.1 (2022), pp. 169–210.

[40] Xin Zhang, Ling Feng, Rongqian Zhu, and H. Eugene Stanley. "Applying temporal network analysis to the venture capital market". In: *The European Physical Journal B* 88.10 (2015), p. 260.