

EXAMINATION OF INSTITUTIONAL INVESTOR NETWORK PATTERNS IN CONTEXT OF
MAJOR CRASHES IN US STOCK MARKETS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS OF
THE MIDDLE EAST TECHNICAL UNIVERSITY
BY

ERSIN DEMİREL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF INFORMATION SYSTEMS

JANUARY 2024

PREVIEW

**EXAMINATION OF INSTITUTIONAL INVESTOR NETWORK PATTERNS IN CONTEXT OF
MAJOR CRASHES IN US STOCK MARKETS**

submitted by **ERSIN DEMİREL** in partial fulfillment of the requirements for the degree of **Master of
Science in Information Systems Department, Middle East Technical University** by,

Prof. Dr. Banu GÜNEL KILIÇ
Dean, **Graduate School of Informatics**

Prof. Dr. Altan KOÇYİĞİT
Head of Department, **Information Systems**

Prof. Dr. Banu GÜNEL KILIÇ
Supervisor, **Information Systems, METU**

Examining Committee Members:

Assoc. Prof. Dr. Çiğdem TURHAN
Software Engineering, Atılım University

Prof. Dr. Banu GÜNEL KILIÇ
Information Systems, METU

Prof. Dr. Nazife BAYKAL
Information Systems, METU

Date: 18.01.2024

PREVIEW

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Ersin DEMİREL

Signature :

ABSTRACT

EXAMINATION OF INSTITUTIONAL INVESTOR NETWORK PATTERNS IN CONTEXT OF MAJOR CRASHES IN US STOCK MARKETS

DEMİREL, Ersin

M.S., Department of Information Systems

Supervisor: Prof. Dr. Banu GÜNEL KILIÇ

January 2024, 82 pages

In the dynamic landscape of the stock market, accredited investors in the U.S. are required to report quarterly holdings in their portfolios to the Securities and Exchange Commission (SEC) using Form 13F. Although these filings disclose holdings of each investor, it is challenging to extract valuable insights because of the complex structure of the market and the natural tendency of investors to keep their strategies discreet.

This thesis presents an innovative approach to process and enrich historical 13F filings into a dynamic bipartite graph of stocks and investors with rich edge and node attributes in order to shed light into investor behaviors in context of major crashes in the U.S. stock markets. The analysis revealed that the clustering coefficient, significant sale and buy counts, late submitted filings and specific motif counts of the generated graph exhibit substantial changes during market crashes. These findings highlight the potential to use 13F filings and network methods for deeper understanding of market dynamics during downturns, and to be used as useful features for market monitoring systems.

Keywords: Financial Analysis, Stock Markets, Stock Market Crashes, Network Analysis, Network Motif Discovery

ÖZ

ABD HİSSE SENEDİ PİYASALARINDAKİ BÜYÜK ÇÖKÜŞLER BAĞLAMINDA KURUMSAL YATIRIMCI AĞI ÖRÜNTÜLERİNİN İNCELENMESİ

DEMİREL, Ersin

Yüksek Lisans, Bilişim Sistemleri Bölümü

Tez Yöneticisi: Prof. Dr. Banu GÜNEL KILIÇ

Ocak 2024, 82 sayfa

Hisse senedi piyasasının dinamik ortamında akredite yatırımcılar, portföylerindeki varlıkları her üç ayda bir Form 13F raporu ile Menkul Kıymetler ve Borsa Komisyonu (SEC) kurumuna bildirmektedir. Her ne kadar 13F raporları yatırımcıların borsadaki varlıklarını açıklasa da, piyasanın karmaşık yapısı ve yatırımcıların stratejilerini gizli tutma yönündeki doğal eğilimleri nedeniyle bu raporlardan değerli içgörüler elde etmek zordur.

Bu tez, ABD hisse senedi piyasalarındaki büyük çöküşler bağlamında yatırımcı davranışlarına ışık tutmak amacıyla, tarihsel 13F belgelerini, zengin uç ve düğüm özelliklerine sahip hisse senetleri ve yatırımcılardan oluşan dinamik iki taraflı bir ağ halinde işlemek ve zenginleştirmek için yenilikçi bir yaklaşım sunmaktadır. Analiz, oluşturulan ağın kümelenme katsayısı, önemli satış ve satın alma sayıları, geç teslim edilen 13F raporları ve spesifik motif sayılarının piyasa çöküşleri sırasında belirgin değişimler gösterdiğini ortaya çıkarmıştır. Bu bulgular, kriz sırasında piyasa dinamiklerinin daha derinlemesine anlaşılması ve piyasa izleme sistemlerinde kullanışlı öznitelikler olarak 13F raporlarının ve ağ analizi yöntemlerinin kullanılma potansiyelini vurgulamaktadır.

Anahtar Kelimeler: Finansal Analiz, Hisse Senedi Piyasaları, Hisse Senedi Piyasa Çöküşleri, Ağ Analizi, Ağ Motif Tespiti

I dedicate this work to my family, who has always supported and encouraged me in my decisions. Freely exploring, learning, and specializing in my career and the subjects I was truly interested in was not something I could do alone.

I hope I made all of you proud.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my social network analysis teacher and advisor, Banu Günel Kılıç. Her guidance has not only made my work possible but also provided me with academic research skills, knowledge, and perspectives that I couldn't find anywhere else.

I am also grateful for the METU Informatics Institute, a supportive and highly research-driven environment where I consistently find answers to all of my questions. As someone from the design background and deeply impressed by the power of network science, I genuinely believe that METU Informatics Institute provides a multidisciplinary ecosystem. I believe that this environment catalyzes exponential growth in innovative ideas that will to shape the future by simply enabling collaboration of curious minds from various domains to create something bigger from themselves.

Thanks for everything.

PREVIEW

TABLE OF CONTENTS

ABSTRACT.....	iv
ÖZ.....	v
DEDICATION.....	vi
ACKNOWLEDGMENTS.....	vii
TABLE OF CONTENTS.....	viii
LIST OF TABLES.....	xi
LIST OF FIGURES.....	xii
LIST OF ABBREVIATIONS.....	xv
CHAPTERS	
1 INTRODUCTION.....	1
1.1 Objectives in the Study.....	1
1.2 Contributions of the Study.....	2
1.3 Organization of the Thesis.....	3
1.4 Scope and Limitations.....	4
2 BACKGROUND AND RELATED WORK.....	5
2.1 Foundations of the Finance and Capital Markets.....	5
2.1.1 Brief Introduction into Finance and Economics.....	6

2.1.2	Capital Markets, Regulations and Disclosures	7
2.1.3	Investment and Its Relevance with Disclosure Monitoring	9
2.1.4	13F Filings, Institutional Investors and Smart Money	11
2.1.5	Bubbles, Bursts and Market Crashes	14
2.2	Data Science in Finance	16
2.2.1	Data Science in Capital Markets	16
2.2.2	Interconnectivity, Networks and Network Analysis Methods	18
2.2.3	Network Analysis in Capital Markets	20
2.2.4	Graph Analysis for Stock Bubbles and Crashes	20
3	METHODOLOGY	23
3.1	Methodology Overview	23
3.2	Tools and Implementation Details	24
3.3	Data Processing for Market Analysis and Bipartite Network Construction	25
3.4	Dynamic Network Analysis and Pattern Discovery in Stock-Investor Relationships	27
4	EXPERIMENTS AND RESULTS	31
4.1	Data Processing, Analysis and Network Construction	31
4.1.1	Description of the Data	31
4.1.2	Data Transformations and Feature Extraction	34
4.1.3	Construction of Bipartite Network	41
4.1.4	Examination of Bubbles and Crashes	44
4.2	Characteristics of the Dynamic Network and its Examination	47
4.2.1	Investor Filing Delays Analysis	48
4.2.2	Significant Portfolio Changes of Investors	51

4.2.3	Network Metrics	53
4.2.4	Network Motif Frequencies	56
5	DISCUSSIONS AND CONCLUSIONS	67
5.1	Discussions	67
5.2	Conclusions	68
5.3	Future Work	70
	REFERENCES	73
	APPENDICES	

PREVIEW

LIST OF TABLES

Table 1	Latest updated 5 forms and their descriptions from official SEC website based on Securities Exchange Act of 1934 Statute	9
Table 2	Source data is a large dataframe of 48.719.087 rows, extracted from historical 13F filings	32
Table 3	SEC API data is used to further enrich information available in 13F filings	33
Table 4	Price Data consists of OHLCV data of available stocks	33
Table 5	S&P500 Price Data also consists of OHLCV data of overall S&P500 companies	34
Table 6	Industry Sectors and Corresponding SIC Code Ranges	38
Table 7	Multigraph is main repository of processed data with all quarters combined	42
Table 8	Summary of Major Market Crashes and Their Key Sectors	48

LIST OF FIGURES

Figure 1	Financial Markets can be classified based on various factors	8
Figure 2	Form 13F Filings are structured documents of portfolio holdings of institutional investors, publicly available in EDGAR database	12
Figure 3	Historical S&P 500 data shows noticeable declines in recent history	15
Figure 4	Data science is an umbrella term encompassing various concepts	17
Figure 5	Graph is a data structure consists of nodes and edges	19
Figure 6	High level overview of the methodology	24
Figure 7	Data processing steps depicted in diagram aims first two objectives	25
Figure 8	Output of the first objective is a dynamic bipartite graph of stocks and investors with rich features	27
Figure 9	Third and fourth objectives of the study is to examine graph structure and change of frequency of motifs over time	28
Figure 10	Motif rules has high explainability and customization	29
Figure 11	While most of filing types are 13F-HR, there are also amendmends and notice reports	32
Figure 12	S&P500 prices data at the time period of utilized 13F filings	34
Figure 13	Filing delays are populated around the 45th day since its due date for submission	35
Figure 14	Holdings per quarter are converted into changes to be displayed as interactions	36
Figure 15	Share changes are categorized based on their values and prevalence in data	36
Figure 16	Investors with 5-50th percentile are inactive, 50-95 are average, and more than 95th are active classified investors	37
Figure 17	Stocks with 5-50th percentile are small, 50-95 are medium, and more than 95th are large classified stocks	38
Figure 18	Sector mentions in 13F filings are distributed as follows	39
Figure 19	Volume and Volatility Category Distributions	40

Figure 20	Quarterly aggregated volume and volatility of S&P500 index	40
Figure 21	Node and edge features of the created bipartite graph of stocks and investors	41
Figure 22	Most frequently interacted 10 investors	43
Figure 23	Most frequently interacted 10 stocks	43
Figure 24	All three crisis have increased volume and volatility, and decreased price with most significant being 2008 crisis	44
Figure 25	Popular tech companies display increase in volume and volatility during DotCom bubble	45
Figure 26	Biggest oil companies show similar hikes during oil crises	46
Figure 27	Biggest oil companies show similar hikes during oil crises	47
Figure 28	Most of the filings are submitted before 45 day delay	48
Figure 29	Filing Delay Category Counts By Industry Over Time	49
Figure 30	All industries have similar delay counts since investors submit all stocks in same filings	50
Figure 31	Hikes in percentage of late filings after market crashes are visible	51
Figure 32	Significant change (Significant Buys and Sales) counts resembles S&P500 volume and volatility graph, highlighting the impact of 2008 crisis	52
Figure 33	Large Buy and Sale counts gets closer during market crashes	52
Figure 34	Hikes in count of large sales are lagged during the Oil Crisis	53
Figure 35	Slight increases in density during crashes are visible	54
Figure 36	During DotCom crisis, software industry average degree visibly dropped	55
Figure 37	During crises, especially 2008, clustering coefficient significantly increases	55
Figure 38	This motif rule focuses on occurrences from inactive to active investors	56
Figure 39	During market downturns, motif counts for corresponding industries visibly increased	57
Figure 40	This motif also filters late submitted filings on top of the first motif	57
Figure 41	Almost all motifs occurred at the late quarters of market crashes	58
Figure 42	This motif captures overlapping sales of active investors	58
Figure 43	No significant changes, although, sales are increasing before market crashes as expected	58

Figure 44	Motif 4 is an inverse version of motif 1 with a changed flow from active to inactive investors	59
Figure 45	Not visible changes, but more counts overall compared to the first motif	59
Figure 46	Motif 5 is an inverse version of motif 2 with a changed flow from active to inactive investors	59
Figure 47	Similar to motif 2, motif 5 counts also has increased in the late periods of 2008 and oil crisis crashes	60
Figure 48	Motif 6 is similar to motif 3, although it counts significant sales of inactive investors	60
Figure 49	Limited number of motif counts because investor activities are based on activity scores, which are calculated depending on sale counts and significance	61
Figure 50	This motif captures any sale and buy edges available in the graph	61
Figure 51	No significant shifts during market downturns, but there is a visible upward trend after late 2012s	61
Figure 52	Motif 8 is similar to motif 7, however, the direction is inversed in this one	62
Figure 53	Similar with the seventh motif, with more counts and volatile increase in counts after 2012s	62
Figure 54	These motifs capture counts of flows of stocks that have high volatility in corresponding quarters	63
Figure 56	These motifs capture counts of flows of stocks that have low trading volume in corresponding quarters	64
Figure 55	Motif 9 and 10 has similar trends since overall market volatility was also increased during these periods.	64
Figure 57	Although varies in counts, motif 11 and 12 has somehow inverted display during 2008 crisis	65

LIST OF ABBREVIATIONS

SEC	Securities Exchange Commission
U.S	United States
S&P500	Standard and Poor's 500
R&D	Research and Development
IPO	Initial Public Offering
EDGAR	Electronic Data Gathering, Analysis, and Retrieval System
XBRL	Extensible Business Reporting Language
Regulation FD	Regulation Fair Disclosure
CAPM	Capital Asset Pricing Model
SMB	Small Minus Big, Market Capitalization Factor
HML	High Minus Low, Value Premium
RMW	Profitability Factor
CMA	Investment Factor
EMH	Efficient Market Hypothesis
ESG	Environmental, Social and Governance
NLP	Natural Language Processing
GraphSAGE	Graph Sample and Aggregate
LSTM	Long-term Short Memory
RNN	Recurrent Neural Network
GRU	Gated Recurring Unit
VAR	Variational Auto Encoder
GAN	Generative Adversarial Network

GCN	Graph Convolutional Networks
GAT	Graph Attention Networks
TGN	Temporal Graph Networks
JODIE	Joint Dynamic User-Item Embeddings
DYREP	Dynamic Learning Representations
CIK	Central Index Key
CUSIP	Committee on Uniform Securities Identification Procedures
API	Application Programming Interface
LLM	Large Language Models
HFT	High-Frequency Trading
PERL	Practical Extraction and Reporting Language
13F-HR	13F Holding Report
13F-NT	13F Notice Report
SIC	Standard Industrial Classification
OHLCV	Open, High, Low, Close, and Volume
NaN	Not a Number
TIAA	Teachers Insurance and Annuity Association of America
IBM	International Business Machines Corp
OPEC	Organization of the Petroleum Exporting Countries
NEC	Not Elsewhere Classified
REIT	Real Estate Investment Trusts
LPPL	Log Periodic Power Law
DJIA	Dow Jones Industrial Average

CHAPTER 1

INTRODUCTION

Due to their diverse, rich, and accessible data, combined with their intrinsic connection to the social sciences, stock markets have consistently captured the attention of investors, economists, and scholars. Ensuring transparency in these markets is critical, both for fair trading and for the protection of individual investors. Over the years, this need for transparency has led to various regulatory requirements, one of the most significant being the SEC Form 13F. Accredited investors use this form to disclose their portfolio holdings in each quarter. Even though the nature of the information presented in 13F filings has a great potential to explain market behavior, extracting meaningful insights from these filings is challenging due to complex structure of market and the natural tendency of investors to keep their strategies discreet.

In this context, this thesis presents a methodology to process and enrich historical 13F filings in order to shed light into investor behaviors within the context of major market events in U.S. stock markets. The process involves the transformation of 13F filings and related stock market data into a dynamic bipartite graph, incorporating rich edge and node attributes. Subsequently, the study delves into the analysis of specific attributes and graph metrics over time, with a concentrated focus on the timelines of market crash events and related industries in U.S. stock markets. Then, network motif discovery approach is applied to examine frequencies of certain network patterns over time, to discover changes of investor behavior during market crashes. In short, study processes and examines historical 13F filings and supplementary data to progressively examine to understand investor motivations and behaviors in relation to market crashes.

The goal of this introduction chapter is to give readers a clear understanding of the motivation behind this research. The introduction chapter starts with the Objectives of the Study section, where main goal and key objectives to reach that goal will be listed. In the Contributions of the Study section, the potential benefits of this study to finance and graph data science fields are discussed. In the Organization of the Thesis section, chosen approach to convey information and the rationale behind it is briefly explained. Finally, the Scope and Limitations section highlights specific focuses within the vast domains of finance and data science, while specifying limitations stemming from the nature of the available data.

1.1 Objectives in the Study

The study utilizes network-centric approaches to better understand the characteristics of the U.S. stock markets during market crises. Knowing the fact that stock markets are highly interconnected ecosys-

tems with numerous large players with serious impact on the overall market or even globally, study examines broad market data as well as industry perspective of stocks, and institutional investor stock holdings.

This main goal can be further divided into the following list of several objectives:

- Transforming the stock holdings data of accredited institutional investors, as disclosed in 13F filings, into a dynamic bipartite network representing stock holding exchanges of investors for each quarter.
- Examination of the overall market and specific major company stocks during stock market crashes of selected industries and sectors.

These initial objectives are expected to pave the way for the following complex objectives:

- Evaluating dynamic network attributes and metrics over time, particularly during market events, to assess changes in the global structure of the bipartite network of stock-investor relations.
- Conducting a thorough analysis of the dynamic bipartite network by examining motif count frequencies each quarter focusing on relevant industries, and comparing their variations during market crashes in U.S. markets.

1.2 Contributions of the Study

Although infrequent, financial crises are significant events with global consequences, often originating from stock market downturns rather than natural disasters or wars. From this perspective, financial crises are events examined from various perspectives.

This study centers on the notion that stock markets constitute a complex and dynamic interrelations of public companies, industries and institutional investors. As a result, the primary contribution of the study to the literature lies in offering an alternative perspective on stock market crashes by examining changes in such interrelations in context of stock market crashes.

Considering the significance of events and the magnitude of the scope, the data to be processed as the ground truth consists of SEC Form 13F filings arising from government-regulated regulations, supplemented by stock market data. In this regard, the thesis also provides alternative approaches for extracting insights from SEC 13F filings.

The study introduces a network-centric approach to the realm of financial market crashes through the analysis of 13F filings. This approach offers an alternative perspective and potentially unique features to be used further in this domain. The findings also indicate the possibility of incorporating rich features for future applications, including the monitoring of financial crises, as well as, prediction or the classification of quarterly positions of investors.

In addition, the thesis provides an in-depth literature review on various related topics such as financial markets, disclosure regulations, filings, graph data science, and stock market crashes. Through this

extensive exploration, it equips the reader with the tools and thought processes to comprehend the interconnections among these subjects.

In this regard, the study aims to provide specific insights by examining investor behaviors in the context of stock market crashes through a network-centric perspective. These insights not only provide finance professionals with an alternative viewpoint on stock market crashes, but also offer investment researchers valuable information to make informed and potentially profitable decisions in the wake of such market downturns.

1.3 Organization of the Thesis

The study commences by establishing an exhaustive background in the fields of finance and data science. Given that the intended readers may have diverse backgrounds, spanning both social science and data science, the upcoming chapter initially provides foundational knowledge on finance and capital markets. It offers general information about financial markets, their fundamentals, regulations, filings, the 13F filings used in the thesis, and background information on financial market crashes. Subsequently, the focus shifts towards the data science perspective, covering the use of data science in finance, the fundamentals of network science, background on the application of network analyses in finance, and relevant examples. Finally, the chapter concludes by presenting related works on the use of graph analysis specifically in stock market crashes.

In the methodology chapter of the study, the efforts required to achieve the objectives presented in this chapter are explained. Each section of the methodology chapter consists of block diagrams explaining steps taken to address each related objective, summaries of the methods to be used in upcoming chapter, accompanied by a rationale for their selection, with references to similar studies.

The "Experiments and Results" chapter serves as the primary section where all the steps outlined in the methodology are elaborated upon in depth, accompanied by relevant figures and commentary on findings. The chapter is divided into two sections. The first section addresses the initial two objectives, centering around the processing and exploration of data. The second section is dedicated to the examination of the processed data, addressing the last two objectives of the study.

First section of "Experiments and Results" chapter begins with an explanation of the data used in the study, presenting descriptive statistics and explanations of its columns within the context of the study. Next subsections details the steps taken to process the data into meaningful features, providing insights into the characteristics of the processed data with accompanying figures illustrating distributions. This section continues to address the first objective, explaining the construction and overall characteristics of the generated graph. Next subsection visualizes volume and volatility data of selected stocks during market crashes to highlight how diverse responses exhibited by each industry during these events.

Moving forward, second section of the chapter addresses the third objective by calculating attributes of the generated graph, such as changes in counts and filing delays, as well as, network metrics like density, average degree, and clustering coefficient for each quarter. This analysis unveils temporal changes, and enables further discussion and potential explanations on how market crashes may influence investor behaviors or vice versa. The concluding subsection of the chapter employs a similar approach to visualize changes, this time focusing on selected graph motifs. Each graph motif is

explained, illustrated and calculated for each quarter to display changes of overall market behavior. Findings are supported with potential interpretations for the observed patterns.

In the Discussions and Conclusions chapter, the discussions section delves into the interpretation of these findings, exploring the utility of the applied approach, discussing its strengths and weaknesses. In the conclusions section, the study concludes its findings by comparing obtained results with the initially outlined objectives and experimental steps taken. Lastly, in the Future Works section of the chapter, potential directions for enhancing the study's utility for regulatory bodies or investment researchers are examined for further improvements in the future.

1.4 Scope and Limitations

Primary data examined is the processed aggregate of historical SEC Form 13F filings falling within the range of 1999 to 2017 years. This limits the scope of the research to this time frame and the market crashes within it, of those occurring in the United States.

In the examining of the market crashes, the study adopts a macro and network-centric perspective. Consequently, the research provides an overall view of the S&P 500, and groupings based on selected industries. As the study is not finance-centric, it solely investigates financial metrics such as price, volume and volatility, refraining from an in-depth analysis of financial metrics. Instead, the study primarily focuses on overall graph metrics, behaviors of industry groups, and changes of investor behaviors as whole. To stay within the defined scope and prevent divergence, the study minimizes emphasis on a entity-level perspective or entity features as much as possible.

Additionally, the study refrains from making predictive assertions and is conducted with an examination approach. The primary reason for this is the low time granularity, and high sparsity of the 13F filings data processed and modeled in the study.

Covering the U.S. stock market in the time span from 1999 to 2017, the examined 13F filings encompass portfolio holdings from over ten thousand unique investors and approximately fourteen thousand distinct stocks. Reflecting the inherent nature of the stock market, both investor and stock interactions follow a long-tail distribution. This implies that a minority of investors and stocks dominate the market with substantial holdings and interactions, while the majority can be regarded as entities with minimal interactions. In addition to that, despite the study's comprehensiveness, filing submissions occur on a quarterly basis, resulting a total of 74 quarters in the time series.

The characteristics of the processed data pose a challenge for the initial modeling approach with graph neural networks, and employing techniques such as graph embedding and link predictions. The nature of the data, involving 74 time frames and over 100 million potential interaction pairs between stocks and investors, leads to suboptimal performance. Consequently, these limitations have prompted the study to adopt a rule-based motif count approach in graph data science, allowing for more explainable findings and preventing potential errors or false claims that could arise from a probabilistic predictive approach.

CHAPTER 2

BACKGROUND AND RELATED WORK

The primary objective of this chapter is to expand upon the concepts mentioned in the preceding chapter, providing comprehensive details and showcasing similar studies from the literature. This approach aims to prepare readers for the subsequent discussions on methodology and experiments.

The concepts and studies presented in this chapter draw parallels from both data science and social sciences perspectives. To ensure that the reading remains on a solid foundation, the chapter is divided into two sections. In the first section, the chapter commences by presenting the foundational principles of finance from a social sciences viewpoint. In this section, capital markets, regulatory bodies, disclosures and their effects are further examined. At the final subsection, directly related studies examining 13F filings and market crashes will be evaluated. It then shifts its focus towards the data science lens in the second section, inspecting data science and computational methods. In final subsection, related studies which specifically apply network-centric tools to analyze financial markets will be examined in depth.

The rationale behind the flow of this chapter is to guide readers through a logical progression towards the main concepts central to this thesis. Even though they are out of the scope of this work, the first section touches upon topics from economics and fundamental theories related to asset pricing, while second one delves into data science methods used in different fields of finance such as portfolio optimization and price prediction because of their prevalence in finance.

In essence, by the end of the "Background and Related Works in Literature" chapter, readers are expected to become familiarized with the concepts and similar studies highlighted in this thesis from both social science and data science perspectives.

2.1 Foundations of the Finance and Capital Markets

This section delves into the foundational concepts revolving around finance and capital markets. It begins with a broad overview of the finance domain, then examines issues related to its specific challenges and nuances. Then it highlights the significance and limitations of the primary data examined in this thesis, notably the 13F filings. It continues with providing related studies in analyzing 13F filings as well as studies addressing market crashes and bursts.

2.1.1 Brief Introduction into Finance and Economics

Finance, economics, accounting along with the law are fundamental and interrelated disciplines that shape the dynamics of our society in global scale [1]. In broad sense, economics, or macro-economics to be more exact, delves into how societies allocate scarce resources among competing uses, examining individual and collective decision-making processes that influence supply, demand, and wealth distribution. Finance, specifically addresses the ways in which individuals, institutions, and markets raise, allocate, and invest capital by determining the value of assets and liabilities. Accounting on the other hand, complements these areas by providing the structured means to record, report, and analyze financial data, ensuring that businesses operate transparently.

Finance, in its essence, concerns the art and science of money management [2]. Its segmented in three primary areas based on its target entity: personal, public, and corporate finance. Personal finance revolves around the financial undertakings of individuals, spanning activities like planning, saving, and borrowing. Public finance, conversely, deals with fiscal strategies, including taxation, budgeting, and debt management, governing how public services are financed. Corporate finance, which is the central focus of this thesis, delves into the financial operations inherent to corporations.

Before evaluating the role and purpose of a corporation, it is essential to recognize the Friedman doctrine, which posits that the primary social responsibility of a business is to increase its profits [3]. There are various paths for businesses to increase profits, from cost reduction and turnover increase to enhancing efficiency and productivity. Naturally, these efforts require certain capabilities, workforce, and material resources. This brings us to the act of 'financing,' the verb form of finance, representing the action of providing funds for an investment [2].

Financing can be viewed as a strategic tool, leveraging the time value of money to utilize anticipated future cash flows for current projects. The principle can be expressed through the formula:

$$FV = PV \times \frac{(1 + r)^n - 1}{r}$$

This equation represents the future value (FV) of a present value (PV), factoring in a given rate of return (r) over a number of equal payments at regular intervals (n).

In a financial ecosystem, while some entities possess a surplus of capital, others might be in need of it. This discrepancy forms the foundation for the 'market for money', where entities with surplus seek investment opportunities and those in need of financing search to secure funds. At the heart of this exchange lies the decision between two primary methods of financing: debt and equity. Debt financing revolves around borrowing funds with the obligation of repayment with interest, while equity financing involves exchanging ownership stakes of the company for capital.

Corporations, guided by their primary responsibility of profit maximization [3], often face the pivotal decision of choosing between debt and equity financing to optimize their operations. A cornerstone in this discourse is the Modigliani-Miller theorem, which suggests that, in a world free from market imperfections, how a firm chooses to finance its operations would not matter [4]. This fundamental theorem sets a benchmark to evaluate external factors that constantly disrupt this ideal scenario.