Networks of Breakthrough Technologies and their Use in Strategic Games for Competitive Advantage

by

Nissia F. Sabri

M.S., Physics, Grenoble Institute of Technology (2006)
M.S., Nuclear Engineering, University of Florida (2007)

Submitted to the Department of System Design and Management in partial fulfillment of the requirements for the degree of

Master of Science in Engineering and Management

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Author .... Signature redacted
Department of System Design and Management
May 20, 2016

Certified by .... Signature redacted ....
Olivier de Weck
Professor of Aeronautics and Astronautics and Engineering Systems
Thesis Supervisor

Certified by ... Signature redacted ....
Alessandro Bonatti
Associate Professor of Applied Economics
Thesis Supervisor

Accepted by .... Signature redacted ....
Patrick Hale
Director, System Design and Management Program

Networks of Breakthrough Technologies and their Use in Strategic Games for Competitive Advantage

by

Nissia F. Sabri

Submitted to the Department of System Design and Management
on May 20, 2016, in partial fulfillment of the
requirements for the degree of
Master of Science in Engineering and Management

Abstract

Breakthrough technologies sustain competitive advantage and are seen as the engine of growth. These technologies can be developed by leveraging internal know-how, but more often they come from an infusion of external technology. The task of screening and selecting innovative technologies to develop or acquire is challenging and relies on various underlying assumptions. This research proposes a systematic framework of analysis that combines network theory and game theory concepts to analyze a set of breakthrough technologies and the companies linked to them, both in the order of $O(100)$. In this framework, breakthrough technologies are represented as a network where nodes represent technologies and links represent dimensions of similarities between these technologies. Network-level metrics provide proxies for estimating the benefit of a node and the cost of a link. The benefit is derived based on the position of the node in the network, and the cost of a link is estimated based on the similarities of technologies it connects. As firms consider a particular target technology, the framework offers a way to calculate the payoff of following a particular path in the network to attain the target from any one of the technologies already in the firm's portfolio. The model provides a recommendation for the best strategy under specific competitive scenarios. Finally, the application of this method is illustrated with various use cases, to analyze strategic decisions made by companies and to explore some that are ongoing. In particular, this analysis looks at hypothetical two-player strategic games in the energy sector, comparing the competitive positions of SolarCity, Siemens and Google to conclude that all three companies have dominant strategies to invest in this sector. The framework was also applied to a strategic game where Google competes with Magic Leap, in the bio-fuel sector and showed a dominant position for Google. The last three scenarios analyzed represent real-world cases, two in the autonomous vehicle domain involving Apple and Toyota and Apple and Tesla and one in the robotics domain involving Toyota and Amazon. The analysis showed the existence of a coordination game in the autonomous vehicle sector where collaboration was beneficial for all parties. Finally, in the robotics case involving the sell-off of Boston Dynamics by Google, the analysis showed that Toyota can leverage a first mover's advantage to create a dominant strategy against Amazon.

Thesis Supervisor: Olivier de Weck
Title: Professor of Aeronautics and Astronautics and Engineering Systems

Thesis Supervisor: Alessandro Bonatti
Title: Associate Professor of Applied Economics

Acknowledgments

I would like to express my gratitude to my supervisors Olivier de Weck and Alessandro Bonatti for introducing me to the topic and their support through the learning process of this master thesis. Furthermore I would like to thank the SDM teaching team and staff who have given me the opportunity to grow personally and professionally. In particular, I would like to thank the outstanding team of teaching assistants; John Helferich, Andrew Walz and Shawn Chang for the support. I would like to thank my loved ones; my husband Julien and my son Ayden for believing in me and for keeping me balanced. Finally I would like to thank my mother Zohra and sister Nefissa, for their unconditional love and for being my inspiration in life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The interest for this research topic was sparked by recent developments in the Internet of Things sector, a sector poised to transform our personal lives and industrial systems. Against this backdrop and the fact that objects and products are increasingly interconnected, many companies have started expanding their technology portfolios. Some have begun entering completely new technology domains. As this trend continues, competition will increase for breakthrough technologies as both the number of firms pursuing a given technology increases, and the number of alternative technologies a single firm considers increases.

This observation led us to wonder how firms select breakthrough technologies to use for creating innovative services and products. The scope of this research focuses on key questions relevant both to companies in the Internet of Things sector as well as to companies outside that sector. In particular, we examine the questions, "How will companies select 'breakthrough' technologies to augment their portfolios?" and, "How will firms position themselves in strategic games against competitors considering the same technology?"

Behind the aforementioned high-level questions lie other fundamental questions we explore in this research: "How do we build the appropriate breakthrough technology landscape?"; "How do we link breakthrough technologies together?"; "How do we represent a firm's position on a technology landscape?"; "How do we define the technology options a firm has?"; and finally, "How do we evaluate the payoffs of different strategies in competitive games?"

The interconnectedness of technologies and thus the interconnectedness of firms that develop, or acquire, them lends itself well to a network representation. Therefore, in this study we have selected network theory as the framework for building the technology landscape and for understanding the links between technologies and firms. Additionally, since competitive forces are strong and strategic moves are important in an increasingly crowded technology space, we selected game theory as the framework for analyzing strategic moves. The result is a model used to uncover underlying links between technologies and firms and allow us to simulate strategic games where multiple firms target the same technology for development or acquisition. The output of the model allows us to understand real-world cases and to predict strategic moves that maximize benefit to a given firm.

In **chapter 2**, we define the meaning of 'breakthrough innovations' and provide an overview of the literature search focused on where innovative technologies originate and how firms select, evaluate and integrate them into their portfolios. We also make a case for why this particular problematic is a strategic game. In the second part of this chapter we provide definitions of the methods used in the remainder of the research. Starting with network theory, we cover mathematical definitions, and the theory's use to formalize the representation of complex natural and engineered systems. Then game theory is introduced as a framework suitable for analyzing most business interactions in which multiple participants seek to maximize their benefit in response to other participants' strategies. Terminology used in **Chapter 5** also is introduced. Finally, we provide a short overview of Natural Language Processing, with a particular focus on the use of the IBM-Watson tool in this research.

Chapter 3 describes the methodology of this research. In particular, it explains the framework proposed, the data samples applied for developing the framework and the use cases selected to illustrate how the framework would apply in practice. The framework proposed represents breakthrough technologies and the firms linked to them as interconnected network layers. The particular technologies related to a firm are identified through a semantic analysis of descriptions of the technologies. Starting from a firm's position in the technology network and given a target technology, this framework allows us to evaluate different strategies to acquire the target technology, particularly under competitive conditions. The second part of this chapter addresses the selection of individual breakthrough technologies to build the overall landscape.

Chapter 4 focuses on the definition of several dimensions of technology inter-relatedness. Specifically, we use Natural Language processing to carry out a semantic analysis of the technologies, from which several measures are extracted: (1) concepts; (2) keywords; (3) taxonomy and (4) entities. Technologies and companies are linked together along each dimension if they share at least one element. The strength of the link is represented by the number of elements in common. In this chapter, we compare the four technology networks and two company networks to each other, to select the best dimensions for network representations. Network-level and node-level measures are extracted to help estimate the cost of moving along a given path in the network, and the benefit of a given target technology. Finally, a trade-space of all possible technology pairs is built to help compare the different options. Nodes and edges are assigned benefits and costs respectively, and these are used in the payoff matrices developed in **Chapter 5**.

In **Chapter 5**, we focus on firms' strategies related to building breakthrough technology portfolios. Applying the network of technologies proposed in **Chapter 4**, we analyze where existing firms' portfolios fall on the network. This is done by highlighting the technology linked to the firms (represented by nodes in the network) and using this technology as a starting point to reach target technologies on the network. The exploration of alternative paths from the source technology to the target technology is based on the cost of the path and the benefit the target technology holds. The combination of these two measures results in the payoff of the chosen strategy. Under competitive conditions the payoff calculation is modified, and can lead to rejection of the previous strategy. The key focus in this section is to analyze the evolution of a firm's network as it runs through a series of games (corresponding to different paths on the network) with the aim of acquiring more technologies.

Finally, **Chapter 6** concludes by proposing other areas of research to expand on the topic.

Chapter 2

Context

2.1 Motivation

2.1.1 Technology and Strategy

Breakthrough Innovations - How are they developed? Where do they originate? Innovations and breakthrough technologies are the engine of growth. Bower coined the original definition of disruptive technology and described it as a an innovation that introduces new features, appealing to new market segments [1]. A disruptive innovation becomes breakthrough strategy if the firm is able to use it to fundamentally disrupt the market's operation, thereby setting new rules [2].

Large technology-based firms are at the forefront of screening and developing breakthrough innovations [3], as these innovations confer an important competitive advantage. These innovative technologies can be developed in collaboration with other firms [3], acquired [4] or developed internally [5].

When developed internally, the firm typically solicits the judgment of internal experts. However, a purely internal view cannot capture the benefit of cross-disciplinary technology diffusion. Thus, other methods have emerged to help enrich internal experts' judgment, and combine it with external data. For example, [6] proposed a method that combines experts' opinions with foretasted returns and risks regarding a given technology. The anticipated risks and returns are derived from forecasted patents citations. This model is called the

Black-Litterman model.

The firm also can use more systematic approaches in selecting new breakthrough technologies when the existing technology portfolio is the starting point for identifying alternative new technology development. This question was explored in [5], where the research was focused on technology opportunity discovery (TOD) to drive new technology development from existing technologies and products, with limited resources. The framework presented consisted of: (a) a structured database containing functional information on existing products and technologies, extracted from a list of patents in different fields; and (b) a logic that links the firm's portfolio to new technology development or applications based on semantic functional similarities.

Recent research on innovation management strategies identified a confluence of technology as the starting point for radical innovation. [7] identified three major drivers of breakthrough innovation: importing ideas from broad networks; creating environments that allow for deep collaboration, and technology-market matching. In this context, technology confluence is defined as the "new combination of previously distinct technologies". Such work at the intersection of different technology domains can lead to significant changes and improvements to existing technologies. But what conditions make this confluence happen?

An appropriate selection breakthrough technologies relies on understanding the larger technological landscape and the position of the firm, and its competitors, within this landscape. Traditionally, economic publications have attributed high importance to a firm's knowledge capital, to measure its strategic position [8] [9]. In particular, the focus has been on so-called essential patents that allow a firm to build products compliant with a set of standards. In their paper, Bekkers and Martinelli [10] proposed a different set of indicators that use instead the firm's position in a network of patent citations.

2.1.2 Technology Integration

Integration of innovative technologies within a firm comes with the need to "infuse" these technologies into the firm's products portfolio. New technologies deliver actual value to the firm only after they are infused successfully into existing systems. In [11], the authors propose a systematic framework to "quantify and assess" the impact of such an integration on the existing systems. One underlying assumption is the technical similarity on a relative scale of the existing systems and the new technology to be integrated. This can be seen as functional similarity, where function is defined as a process acting on an operand to change its state [12] [13]. These fundamental processes and operands are defined as (Storage, Transport, Transform, Store, Exchange, Control) and (Matter, Energy, Information) respectively in [14]. Do proposes an extension of the objects to include (Living organisms and Money) [15] . For example, technologies that rely on the same process or act on the same operand are more similar than those that have different processes and operands. This is where linking technologies along a measure of functional similarity can help ensure the firm is considering technology options that will, indeed, be more easily integrated into its existing portfolio.

2.1.3 Technology Options

Whether technology is developed internally or acquired, it is an investment. Thus, all possible options need to be evaluated to ensure the best are selected.

Selecting a technology to develop or acquire depends on the technological capability of the firm, but it also depends on the strategic positioning the firm is targeting for the future. Different methods for identifying critical technologies at an industry level have been described in the literature. For example, [16] combines experts' judgment and system dynamics models to recommend future technology investments in the Chinese ICT industry.

On one end of the technology options spectrum there are publications that emphasize technology domains and thus technical capability. For instance, the think tank 'The Institute for the Future' proposed a map based on combining high level technology domains, to explore potential future technologies at the intersections of different fields [17]. This is shown in Figure 2-1.

At the other end of the spectrum, there is the approach based on competitive intelligence through mapping firms' direct investments and acquisitions, such as the mapping proposed by the commercial service [18] and shown in Figure 2-2.

These conditions lend themselves well to strategic games in which firms represent players evaluating different options.

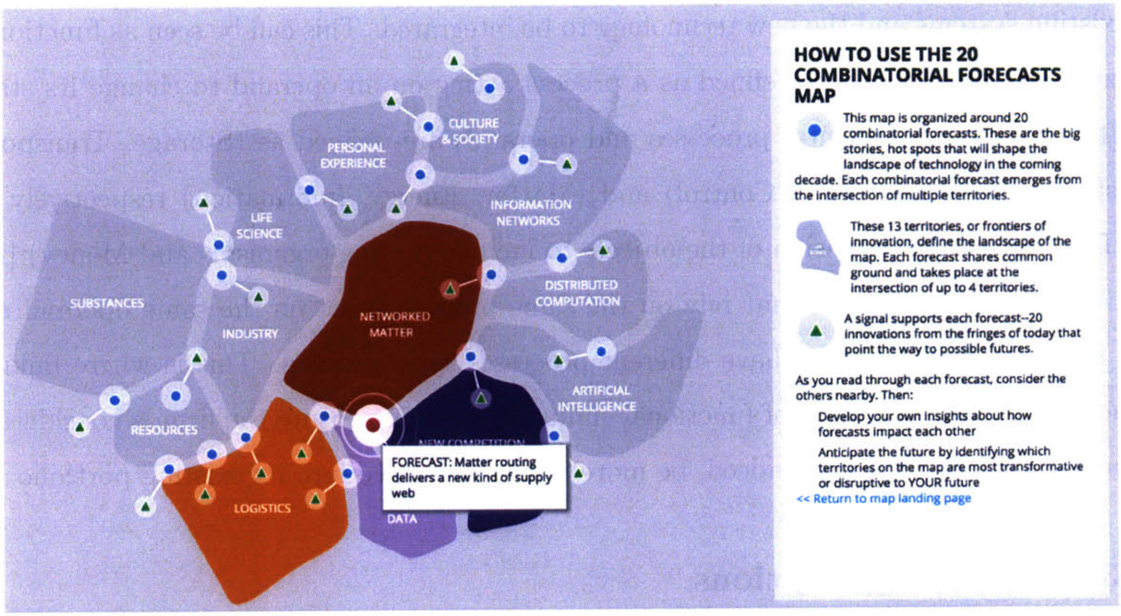

Figure 2-1: Institute for the Future - 20 Combinatorial Forecasts Map

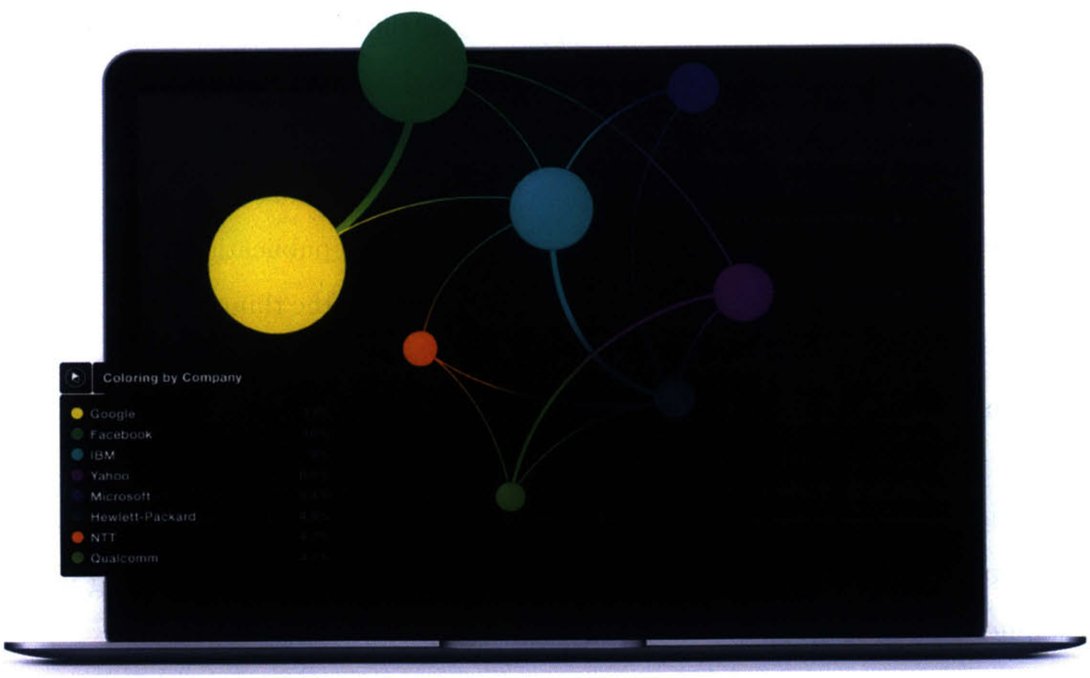

Figure 2-2: Quid - Competitive Intelligence Map

2.1.4 Strategic Games for Technology Investments

Various research papers have looked at different aspects of such a strategic game. For example, [4] discussed how to quantify the option value of technology investments. While Net Present Value (NPV) approaches help select investments that have expected returns exceeding market returns of a similar risk level. NPV and DCF assume a one-time decision, in particular with technology portfolios. However, decisions are dynamic and managers can re-adjust technologies in their portfolios based on market developments and the probability of successful integration of a given technology. This is why the framework of real options proposed in [4] is well-suited, as technology investments inherently include risk as well as growth opportunities.

In his paper, Smit answers two questions: (1) How much is a strategic option worth?; and (2) How does one analyze strategic options in a dynamic, competitive environment?. In this paper real-options and game theory are used to derive the best technology investment. Further, the author argues that "strategic initiatives can no longer be looked at as standalone investments, but rather as links in a chain of interrelated decisions." In [19] the authors propose a mathematical model to examine investment games between two firms under asymmetric information and uncertain revenue flow.

Finally, while most technology acquisitions aim at integrating the novel technology into an existing portfolio of products, in some cases access to a novel technology can be used strategically to block competitors. This is why technology investment can be seen as strategic games.

2.2 Definitions

2.2.1 Network Theory

General view

Networks as a representation of natural systems. Mitchell discussed in her paper [20], how networks relate to complex systems. She argues that complex systems are represented by systems of systems and one representation is through networks. For example, nodes in a network represent smaller, less complex systems from a -functional point of view- that form larger systems through interactions represented by the links.

Some fundamental areas of research on networks try to answer questions regarding appropriate topological measures for characterizing networks, and to characterize propagation in the network of information, failure and so forth. Finally, the most relevant aspect of this research for our analysis is related to how properties of the network can be used to reach a particular node in the network.

Networks representing engineered systems. Network science is characterized by its methodology of bringing together multiple disciplines from social sciences to biology. It has evolved around a set of methods and tools that help represent both natural systems and engineered systems, using similar mathematical formalism. Due to its multi-disciplinary nature, network theory works well for use in research that aims at better understanding connections between breakthrough technologies and firms' strategic advantages in acquiring or developing them. Additionally, network science is data driven and relies on representation of empirical linkages. In fact, its focus on empirical data distinguishes network science from graph theory.

Networks and graphs are used interchangeably but what distinguishes them is that networks (nodes and links) represent real systems while graphs (vertex and edges) are the mathematical or abstract representations of these real systems.

Networks are defined by the entities they represent, referred to as nodes, with the interactions or connections between these entities referred to as links. The value of network representation of complex systems depends on what nodes are represented and what the links connecting them mean [21].

Graphs are either directed or undirected indicating reversibility or reciprocity between two nodes. For some networks, the roles of the source node and the target node are not interchangeable and for others they can be. If, for all nodes in a given network the directionality is represented, the graph is directed. If not all nodes in a graph are directed, the graph itself is undirected.

Types of Networks

In linked [22], Barbasi and Watts describe three general networks: random, small-world and scale free.

1. Random Networks. In random networks, each pair of nodes is linked with a uniform probability. Therefore, there is no strong clustering in random networks and their degree distribution is a Gaussian for large numbers of nodes. In contrast, real-world networks exhibit a power law degree distributions of the form

$$P(k) = k^{-gamma} \tag{2.1}$$

2. Small-world Networks. In small-world networks, each node has k neighbors, for each link there is an associated probability p to be instead connected to a random node. If p=0, the network is the initial regular network, for p=1 it is a completely random network and for p small there are numerous local connections and a few long distance connections.

3. Scale-free Networks. In this network, nodes have preferential attachments and the degree distribution follows a power law. In contrast to Gaussian degree distributions that tend to have a cutoff value where the distribution reaches zero, the power law degree distributions have no cutoffs. In scale-free networks, there are few hubs of small numbers of nodes with a large number of links connected to the rest of the network, where there are more nodes with fewer links.

The network growth model proposed by [23], starts from a network with a small number of nodes and adds a node at each time step that is connected to m nodes in the network. The probability of that node to be connected to node i is proportional to the degree of node i in the network. Finally, according to [23] most real-world scale-free networks follow the power law $P(k) = k^{-gamma}$, with a gamma coefficient between 2 and 3.

There are a number of assumptions in the networks characteristics discussed above: -

- do not account for appearing and disappearing links
- do not account for links weights

- do not account for different dimensions of relatedness
- There are no costs associated in creating links
- Assume that preferential attachment is proportional to the degree of the node only.

When considering networks with different characteristics from those listed in these assumptions, we expect the degree distribution to be different from a power law.

Networks Characterization

Characterizing real-world networks is important to understanding the complex system they represent. Traditionally, network characterization was motivated by understanding the resilience of networks in case of cascading failures, areas of vulnerabilities or propagation within the network. Mostly, this kind of characterization is used to design better systems or to improve processes, such as drug delivery. In the context of this research, the characterization of the network of breakthrough technologies is used to inform decisions. Important structural properties of the networks are discussed in the following sub-sections:

1. Distance Represents the length of the path between two nodes. These paths are non-unique and even algorithms that calculate the shortest path produce non-unique paths that all have the same length, but not necessarily the same number of links and nodes along the path.

2. Weighted paths, diameter and shortest path This is relevant in this research as the cost of moving from one technology to the next is measured by the path length in the network, with the shortest weighted path being the most efficient path to move from one node to the next.

3. Community Communities, or clusters, are sub-networks with dense connections between the nodes. At a node level, this means a node that belongs to a community is much more likely to have connections to nodes within the community than to nodes outside the community. Guimera [24], defined different roles for nodes within a community depending on their links to other nodes and proposed the following definitions for hubs of nodes with many connections within their community:

- ultra-peripheral nodes where all links are within the nodes' module
- peripheral nodes where most links are within the nodes' module
- non-hub connectors nodes where many links connect to other modules
- non-hub kinless nodes where links are homogeneously distributed among all modules

In this paper Guimera also defines the roles for sub-hubs: -

- provincial: the vast majority of nodes links are within the nodes module
- connector: the node is both a hub in its module and has many links to most other modules
- kinless the nodes links are homogeneously distributed among all modules. While the authors applied these definitions to biological networks, they concluded the role of the node may be a better measure of its importance in the network than its degree.

This is relevant for this research as we would like to identify the important technologies in the network and compare their importance to inform decisions.

4. Networks Dynamics Network dynamics represent changes in the network's topological structure or in the exchanges that occur within the network. In addition to understanding the static structure of nodes, to characterize network dynamics we need to understand how individual nodes are linked.

Two examples of networks dynamics are biological systems and information processing. For the purpose of this study, the networks dynamics that can be represented are:

- Changing links between technologies as firms invest or divest, creating new links
- Exchange of capability between two technologies that are linked - either through knowledge, process, or system.

Mathematical view

In network theory a graph G=(V,E) is a set of vertices V=(v1,v2,...vn) and edges E=(e1,e2,e3...em).

1. Adjacency matrix The adjacency matrix **A** is an N by N matrix, in which the columns and the rows represent all nodes in the network, and the value at the intersection of each row and column represents either the existence of a link ($A_{ij} = 1$) or absence of a link ($A_{ij} = 0$) in unweighted networks. In weighted networks, the values represent the weight of the link between two nodes.

$$A = \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix}$$

2. Measures at the Node Level Degree represents the number of links an individual node has k_i; for directed networks the total degree of a node $k_i = k_i^{in} + k_i^{out}$. The degree can be calculated from the adjacency matrix as the sum of the column or the row representing the node.

$$< k_i >= \sum_{i=1}^{n} A_{ij} = \sum_{j=1}^{n} A_{ij} \tag{2.2}$$

For directed networks, the convention is that the incoming and outgoing degrees are calculated as follows:

$$< k_i^{in} >= \sum_{i=1}^{n} A_{ij}, < k_i^{out} >= \sum_{i=1}^{n} A_{ji} \tag{2.3}$$

Clustering . Clustering represents the average probability that two neighbors of a node are also neighbors of each other. The mathematical representation is as follows: for a node i with a degree k_i, the local clustering coefficient is

$$C_i = \frac{2L_i}{k_i(k_i - 1)} \tag{2.4}$$

In this equation, k_i represents the neighbors of node i, L_i represents the number of links between these neighbors. $C_i = 0$ represents the absence of links between neighbors while $C_i = 1$ represents the fact that all node i neighbors are connected. Since C_i is between 0 and 1, it represents the probability that two neighbors of a node are linked.

Centrality measures

(1) Degree centrality measures the nodes with the highest degree [25].

(2) Closeness centrality is a measure related to the length of the paths from a node to all other nodes in the network, and is calculated as the inverse total length.

(3) Betweenness centrality is calculated in two steps: first, the shortest paths in the network are calculated; then the number of them that pass through a given node is calculated. This results in the betweeness centrality measure for a node.

3. Measures at the Network Level Links It relates to the total number of links in the network as follows. Note that the 1/2 is added to account for the double counting of a single link from both nodes it connects: Total number of links in undirected graphs:

$$L = \frac{1}{2}\sum_{i=1}^{n} k_i \tag{2.5}$$

Total number of links in directed graphs:

$$L = \sum_{i=1}^{n} k_i^{in} = \sum_{i=1}^{n} k_i^{out} \tag{2.6}$$

Average Degree . The average degree for an undirected network is

$$< k >= \frac{1}{N}\sum_{i=1}^{n} k_i = \frac{2L}{N} \tag{2.7}$$

and for a directed graphs

$$< k^{in} >= \frac{1}{N}\sum_{i=1}^{n} k_i^{in} =< k^{out} >= \frac{1}{N}\sum_{i=1}^{n} k_i^{in} = \frac{L}{N} \tag{2.8}$$

Degree Distribution . Degree distribution represents the distribution of frequencies of all degrees over the nodes in the network. It represents the probability p_i that node i, taken randomly, has a degree k. By definition, the sum of all probabilities sums to 1. Once the histogram is constructed, representing the number N_k, the number of nodes with degree k, the probability distribution of the nodes degrees is obtained by normalizing N_k over the total number of nodes N.

$$p_k = \frac{N_k}{N}, \qquad \text{for } k = 1..inf \tag{2.9}$$

Deriving an equation for the probability degree distribution allows us to calculate the average degree of a network, which can be calculated as follows:

$$< k >= \sum_{i=1}^{n} kp_i \tag{2.10}$$

Distance . Distance in networks is represented by the path length between nodes. One of the most important path length calculations is to derive the shortest path between nodes. In an unweighted network, the shortest path is the one with the least number of links between nodes i and j. In undirected graphs the shortest path length is the same between the source to target and target to source, while in directed graphs this can be different. The shortest path is not necessarily unique, as there could be different alternative paths between the source and the target, all with the same path length.

For a given distance d_{ij} between nodes i and j and an adjacency matrix described above A with elements A_{ij}, the number of shortest paths with d links is N_{ij}^d and can be theoretically be calculated as follows:

$$N_{ij}^d = \sum_{k=1}^{N} A_{ik}A_{kj} = A_{ij}^2 \tag{2.11}$$

The diameter of the network is the maximum shortest path from the set of shortest paths between any pair of nodes.

The average path length is simply the average of all distances between any two nodes in the network.

Connectedness . A connected network means that all pairs of nodes are connected, in contrast to a disconnected network that has at least one pair of non-connected nodes. A network can have multiple connected components that can be connected through a link called a bridge. In fact, disjoint intra-connected components in a network can be represented by a block diagonal matrix with zeros off the diagonal blocks. The non-zero elements off the diagonal represent the links between connected components. Figure 2-3 shows an example from [26].

Clustering coefficient . The average clustering coefficient in the network is represented by the following equation and shows the degree of clustering of the whole network. This is defined for undirected networks but can also be generalized [21] to directed and weighted networks.

$$< C >= \frac{1}{N}\sum_{i=1}^{N} C_i \tag{2.12}$$

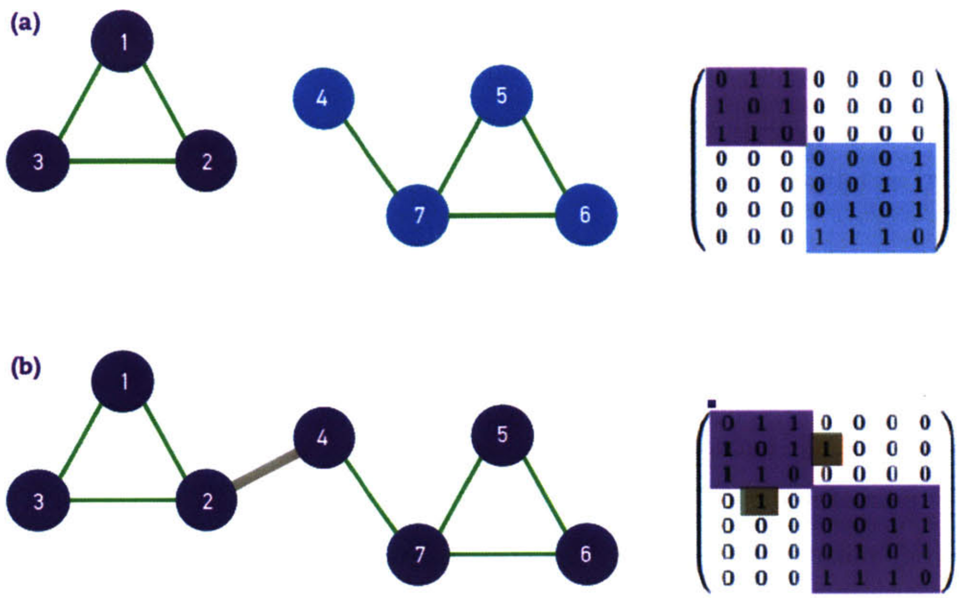

Figure 2-3: Adjacency Cluser

4. A special type of network: A bipartite network is constructed from two disjoints sets of nodes N1 and N2. This network represents the links between nodes in N1 and nodes in N2 without the intra-network links within N1 and N2. There are two projections that can be generated from such a network; showing connections between nodes in N1 if they connect to the same node in N2, and vice-versa for the projection in N1. Figures 2-4, 2-5 and 2-6 from [26] illustrate such a network.

Similarly, we can define multipartite networks that connect more than two disjoint groups of nodes. Additionally, there are references [27] that define other network measures, called network indices, that are not described here.

2.2.2 Game theory

Strategic games are everywhere in social interactions, sports and businesses. But they can be particularly useful for the following: (1) explanation, (2) prediction and (3) prescription.

1. Explanation: Relates to how game theory can help us understand unfolding events

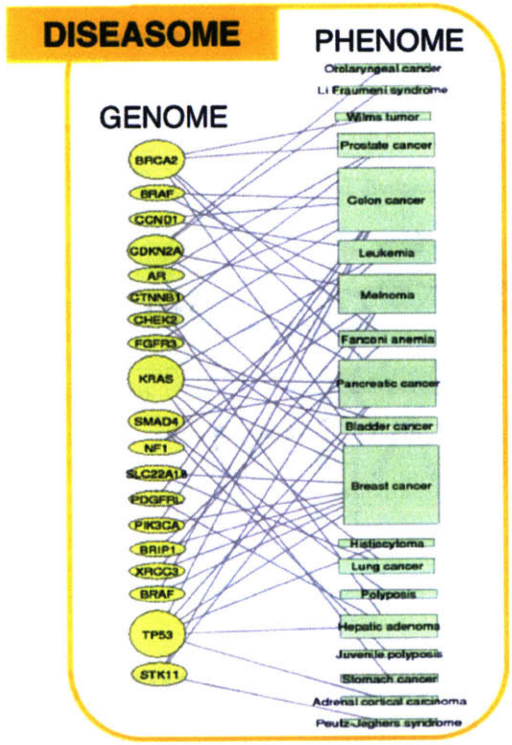

Figure 2-4: Bipartite network example from [26]

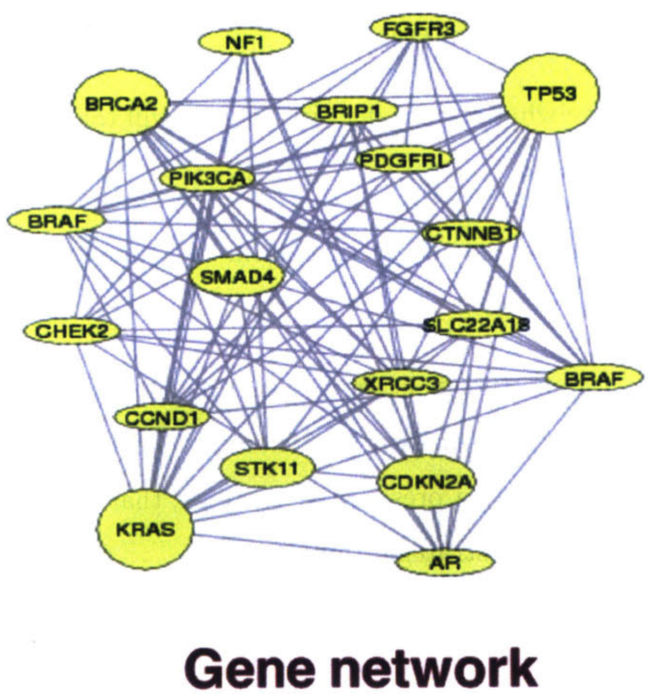

Figure 2-5: Bipartite network example from [26]

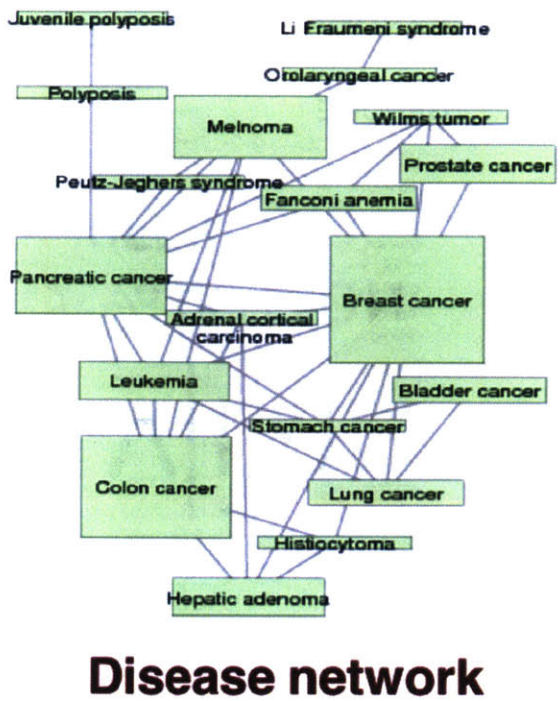

Figure 2-6: Bipartite network example from [26]

when multiple decisions makers interact.

2. Prediction: Helps foresee what actions decision-makers will take based on the situation and the potential outcomes.

3. Prescription: Outline and advise one participant in the game on following the strategy with the best expected outcome.

One important aspect of game theory is the assumption that players are rational, know the game's rules and are acting in their own best interest. There is always a risk in such games, especially when used to apply a prescriptive strategy, that one of the participants is 'clueless' and/or did not go through strategic thinking.

Additionally, in many interactions there is commitment and private information. This means that once participants have made a commitment by engaging with the opponent, the two participants are tied to each other independently of the larger group. Similarly, once an interaction between a few participants has started, they would know more about each other.

Therefore, when considering strategic decisions in which a few participants are important, committed or tied to each other and share private information, these participants become significant in this relationship. This is where the interaction becomes a strategic game.

Games and Strategic Decisions

A strategic game is defined by a cross-effect of the actions of the participants. For a strategic decision to become a game, participants must be mutually aware of the cross-effects of their actions. The rationale of strategic games implies that a participant knows that the actions of the opponent will affect him/her and therefore can react to the opponent's actions, or make his/her own actions to alter the opponent's actions. This distinguishes strategic games from games of chance, or competitions that rely on the skills of the two opponents.

Strategic games are most prominent for analyzing strategic decisions between two opponents. Traditionally, game theory was not used to analyze interactions of a large number of participants. However, even for these situations it turns out to be strategic games between a small number of participants.

Types of Games

The following concepts and definitions are extracted from [28]. The summaries in this section are meant to help familiarize the reader with key terms used in the following chapters but do not necessarily capture all nuances presented in the book.

Sequential and Simultaneous games

> Definition: In sequential games a participant has to consider how the opponent will react if he/she proceeds with a given move. The move in the present is tied to future consequences. In simultaneous games the participant must predict what the opponent will do in the present. The same thinking happens concurrently on the opponent's side.

Conflict and or commonality

Definition: Zero-sum games or constant-sum games arise when opponents must divide gains in a way that one receives more than the other, or one receives everything and the other nothing. Sports games provide an example of this. However, in most economic interactions the games are not constant-sum because participants can generate more benefit combined. Conflict and cooperation often are in tension as they both can be used in a game, with the threat of conflict applied to force opponents to change position. When there are more than two participants, a single game can see cooperation between some participants at the expense of others. War alliances are an example of this.

Repeated games and changing opponents

Definition: A game can be played once, or multiple times, with the same or changing opponents. When games are played repeatedly, strategic thinking must be carried to the outcomes of future moves in the game. In one shot games the opponents do not have to worry about building a reputation, and are less likely to compromise as they seek to maximize their profit from a single encounter. In repeated games, reputation matters, whether with the same opponent or a different one. Therefore, over time, participants seek to build reputations that will endure through the repeated games. They also may be more likely to compromise on a given game, or to choose moves that strengthen the reputation they want to reinforce. Punishing actions, or taking turns, is more likely in repeated games.

Information

Definition: Games of perfect information are rare. In most games there are two types of uncertainties: strategic uncertainties related to past and present decisions of the opponent, or uncertainties related to external factors. In simultaneous games, strategic uncertainty is high but can become higher for a participant who has less information than the other. This is asymmetric information. In some situations, one participant may reveal information to other participants to affect their actions - this is called **signaling**. Signaling requires tangible proof of information revealed, otherwise any participant could 'reveal' false information to manipulate the game. For example, in the R&D world, announcement that a company has invested in building new infrastructure is a signal that is stronger in deterring competition than simply announcing the company intends to pursue some R&D. On the other hand, opponents can test how committed the participant is to something. This is screening, with the intent of filtering relevant and true information from unsupported announcements.

Fixed Rules

Definition: In sports games the rules are fixed and known, and all participants know everyone else is playing under the same rules. In other matters, such as business and life, rules are set by some players and can be manipulated by others. In situations in which only one participant sets the rules, that participant must be credible. Inversely, some participants can make their opponent's threats less credible.

Cooperation Enforceability

Definition: Most real-world games include conflict and cooperation. In games in which participants can meet and negotiate agreements, the game reaches an equilibrium when participants have made a choice and there is no better option for any one of them. Once agreements are negotiated, parties can abide by the rules agreed upon or implement their own actions in private - which may or may not be possible to monitor. The distinction between games in which negotiated agreements can be enforced and those in which they cannot, is important. Cooperative games are those in which agreements are enforceable, whereas those in which agreements are not enforceable are called non-cooperative games.

Table 2.1 summarizes the type of games considered in this research.

Table 2.1: Types of Games Selected

Type of games	Option selected
Timing	Simultaneous
Payoff	Constant-sum
Repeatability	Once, same opponent
Information	Same, private
Rules	Manipulable
Enforceability	Non-enforceable

Terminology

Strategies

Definition: Strategies represent the different options or choices a player has in a game to respond to moves of other players in the game.

Payoffs

Definition: Payoffs are the quantified outcome of a choice in a game, taking into account the cost and the benefit of the move. The payoff number captures all of what the player cares about. Payoffs from random outcomes are the weighted average of payoffs from the possible outcomes, with the weights being the probabilities of the considered outcomes to occur.

Rationality

Definition: We assume all players carry out the correct calculation and will follow the strategy that leads to the best outcome. It is important to distinguish rationality from selfishness. The payoff to be maximized by the rational player includes everything the player cares about, including the well-being or the benefit to other players in the game. The important parameter is calculating one's own payoff, and estimating other players' payoffs by understanding their value systems.

Rules

Definition: Players must have a basic knowledge of the rules. For any game, the rules consist of the following: - (1) The players - (2) The strategies/options each player has - (3) The payoffs that correspond to all possible scenarios - (4) The assumption each player is rationale and aims to maximize his/her own benefit.

Equilibrium

Definition: Equilibrium is reached when each player chooses his/her best strategy in response to other player's strategies. Equilibrium for simple games is easy to achieve but as a game becomes more complex and the number of players increases, computer programs can be very helpful.

In particular, an open source project [29] offers a set of tools to compute more complex games.

2.2.3 Natural Language Processing

This research employs a Natural Language Processing tool from the IBM - Watson suite called AlchemyAPI [30]. While the scope of this research does not include an analysis of the fundamentals of Natural Language Processing, in this section we provide basic definitions of the subset of tools used.

General Definition

Natural language processing (NLP) sits at the intersection of computer science, artificial intelligence, and computational linguistics. At its core, the discipline is concerned with making human language, with its nuances and variations, understandable by computers. This would improve human–computer interaction, in particular the ability to make the complex tasks of translating commands to computers in an non-programming environment. Simply put, the practical use of NLP is in enabling computer programs to read natural language text written by humans and understand its meaning. Of course, this goes beyond a simple aggregation of keywords that appear in the text, instead targeting the conceptual meaning of written words. NLP can be decomposed into : "natural language understanding, enabling computers to derive meaning from human or natural language input; and others involve natural language generation [31]."

AlchemyAPI

AlchemyAPI uses NLP algorithms to analyze input submitted as full texts or URLs, and tags the extracted information. The tagged information is categorized in different groups: entities; keywords; taxonomy; concepts; document sentiment; targeted sentiment; document emotions; relations; language; Title; Author; Text, feeds and micro formats. The categories relevant to this research, and used in the following chapters, are described below:

Concepts

> The AlchemyAPI concept tagging feature can make high-level abstractions from texts. The algorithm understands how concepts relate to each other and can list concepts even if they do not appear in the text. For example, if an article mentions 'CERN' and the 'Higgs boson', the algorithm will tag 'Large Hadron Collider' as a concept even if the term is not mentioned explicitly on the page. Another example would be tagging "Automotive Industry" from the sentence, "My favorite brands are BMW, Ferrari, and Porsche." Figure 2-7 shows an example of concepts extracted from the tool.
>
> These concepts are generated from pre-existing databases such as "Linked Data" [32]. Figure 2-8 shows a conceptual network of how data is linked in this database. "Linked Data is a method of exposing, sharing, and connecting data on the Web via dereferenceable URLs. Linked Data aims to extend the Web with a data commons by publishing various open datasets as RDF on the Web and by setting RDF links between data items from different data sources. The Linked

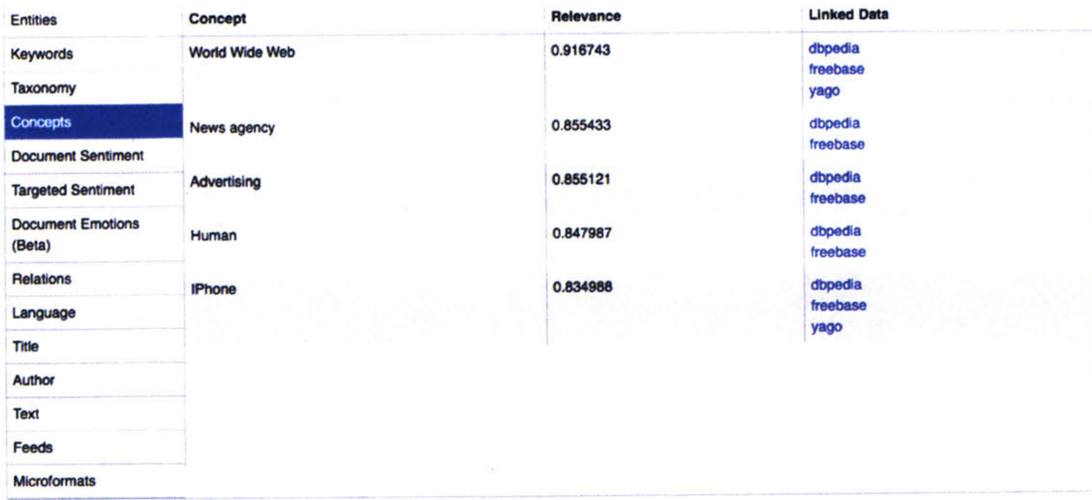

Figure 2-7: AlchemyAPI - Concepts tagging

Data cloud currently consists of over 7.4 billion RDF triples, interlinked by 142+ million RDF links.".

Keywords

Keywords represent the important topics in the text content provided to AlchemyAPI. The output is a list of keywords from the text and their ranking, based on statistical occurrence. The algorithm also can provide sentiment analysis for each of the extracted keywords, although this feature is not employed in this analysis. Figure 2-9 shows an example of keywords extracted from the tool.

Taxonomy

The output of the taxonomy analysis provides classification of the most likely topic in which the described content falls. This hierarchical taxonomy can go as deep as five levels, allowing a more granular categorization of the content analyzed. For example, a text describing personal lending practices can be classified into the following sub-topics:

- /finance/personal finance/lending/credit cards /finance/personal

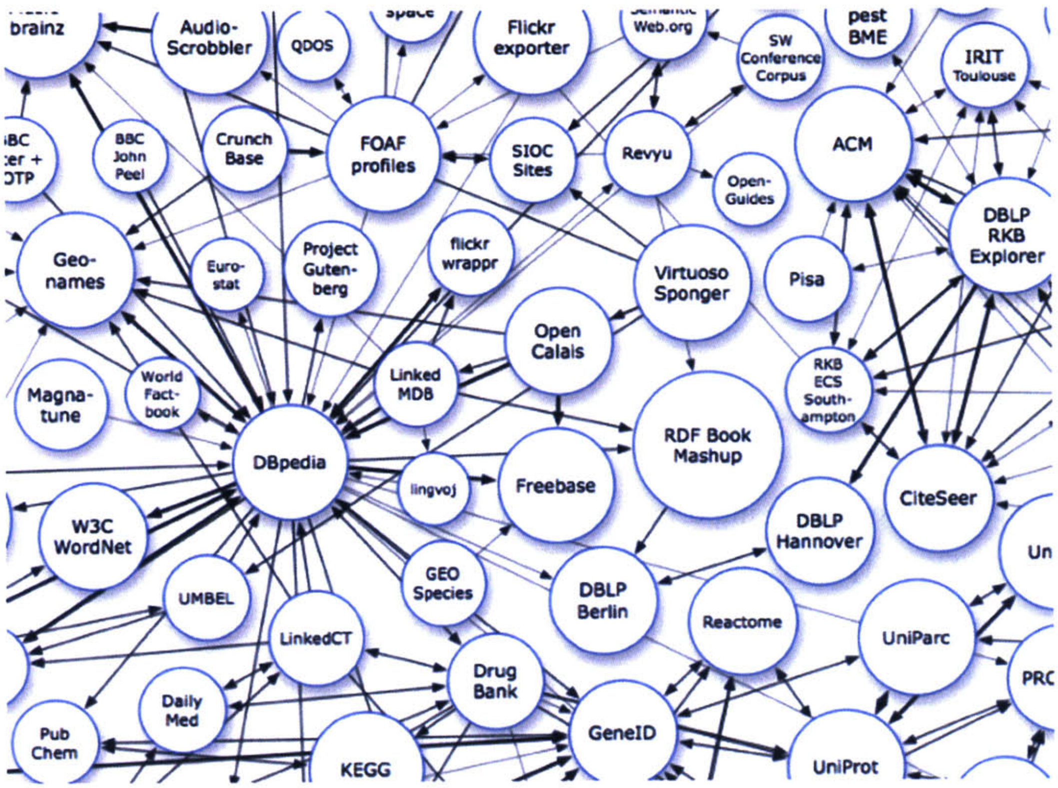

Figure 2-8: AlchemyAPI - Linked Data

Figure 2-9: AlchemyAPI - Keywords tagging

finance/lending/home financing /finance/personal finance/lending/personal loans /finance/personal finance/lending/student loans /finance/personal finance/lending/vehicle financing Figure 2-10 shows an example of taxonomies extracted from the tool.

Entities

The last feature used in this analysis targets entities such as persons, places and organizations. Entity extraction is most often used as a starting point for employing natural language processing techniques. In particular, this feature is of interest in this analysis for extracting companies linked to a given technology. The link is not necessarily direct, since a company would be tagged if it is mentioned in text describing the technology. More precisely, it is used to establish a link within a socio-technical environment between technologies and companies. Figure 2-11 shows an example of entities extracted from the tool.

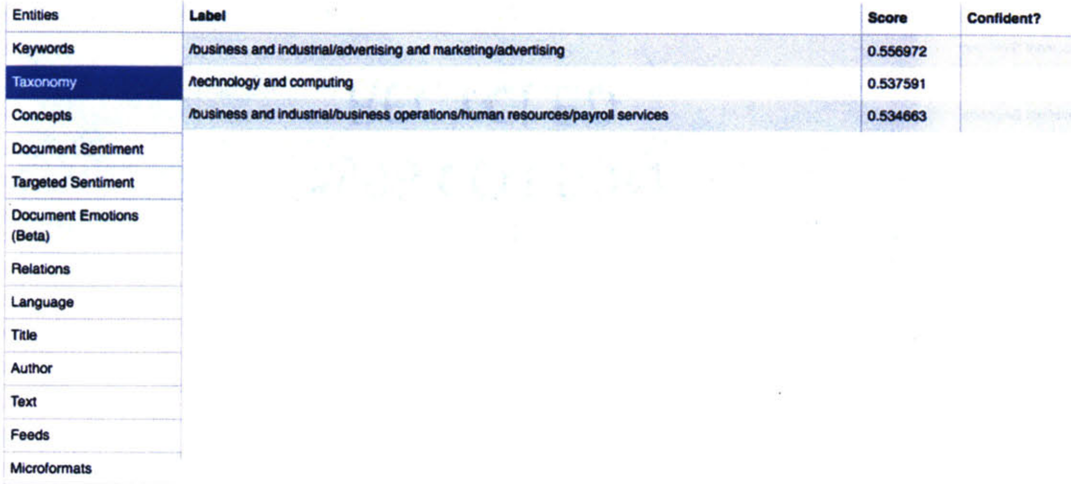

Figure 2-10: AlchemyAPI - Taxonomy tagging

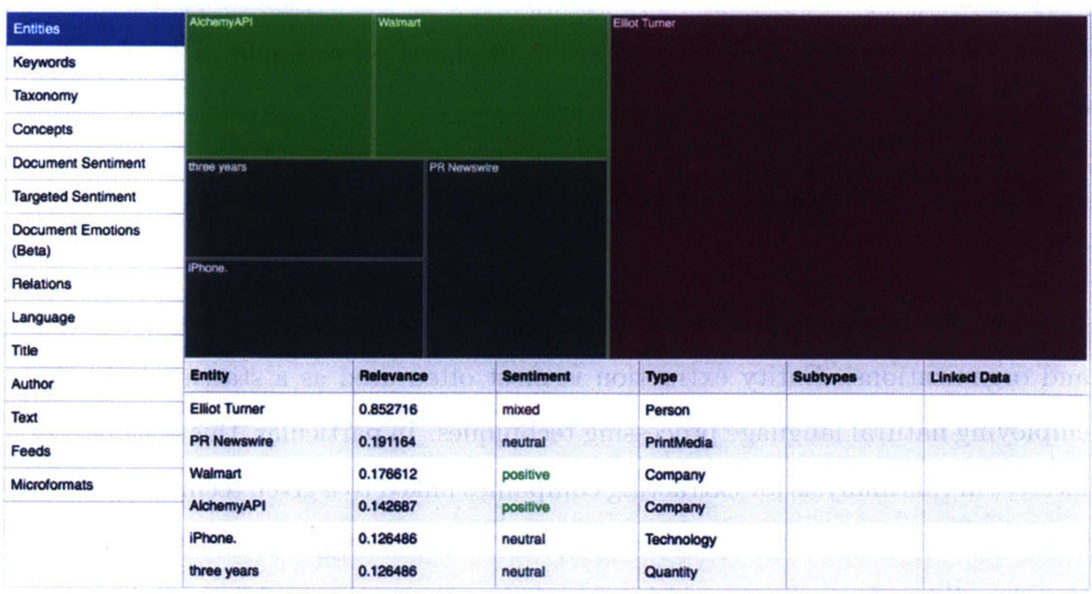

Figure 2-11: AlchemyAPI - Entities tagging

Note that the pictures of the AlchemyAPI shown above illustrate the web interface of the application. In this research we have used the service through a Python API.

Chapter 3

Methodology

3.1 Framework

The framework proposed in this research draws from two disciplines - network theory and game theory - to enable the construction of a technology landscape in which strategic games between firms can be represented and analyzed. The major advantage of this framework, compared to traditional technology portfolio strategies described in **Chapter 2**, is that it builds a network of technologies that connect to each other. And to the portfolio of a given firm. Further, this framework enables analysis and comparison of different strategies that can be used to reach the target technology. It does so by evaluating all the options the firm has given its initial position in the network, and the payoffs of given paths and targets under different competitive scenarios.

The main steps of building the framework are described below and illustrated in Figure 3-1:

1- Build a network of technologies and companies

2- Define linkages between the nodes based on similarities

3- Represent firms positions on the network

4- Evaluate benefits and costs of nodes and edges on the network

5- Evaluate the payoffs of different strategic games in the network

6- Apply to use cases

The following six sub-sections describe the steps listed above in greater detail.

Step 1: Build a Network of Technologies and Companies

To be able to represent different strategic games on a technology network, the network itself must be relevant to the firms of interest, and representative of technologies across-domains. In addition, we limit the list of breakthrough technologies to those representative of a system or a process capable of delivering value on its own. This criteria is important, as fundamental research offers many breakthrough innovations but does not play a key role in firms' strategic decisions. **Therefore the technology list used in building the network needs to be relevant, diverse and representative of mature systems or processes***

In the literature, there is no single comprehensive list of breakthrough technologies. Instead, there are different publications that every year keep track of novel and impactful technologies such as, the MIT Technology Review, World Economic Forum's reports and the OECD database. Other approaches include compilations from patent databases.

The advantage of using technology publications is that they offer a multidisciplinary view and select mature systems and processes. Table 3.1 below shows the different publications considered to build the list of breakthrough technologies.

Table 3.1: Publication Sources of Breakthrough Technologies

Source	Data description	Number	Time span	Pros	Cons
MIT technology review	Generic description	150	2001-2016	Cross-disciplinary, stand-alone system	Not comprehensive, selection process is unknown

World Economic Forum	Generic description	30	2013-2015	Global view, cross-disciplinary, functioning system	Not comprehensive
OECD	Country Statistics	R&D research	TBD	Product, but necessarily technology breakthrough	Comprehensive
Wikipedia	Not vetted	TBD (high)	Unknown	Cross-disciplinary list and functioning systems	Pedigree not validated

Other approaches investigated are: selecting a list of firms first and building the technology landscape around them, or selecting high-level technology domains and mining patent databases for recent technologies. Although these approaches are not used in this analysis the sources of data are listed in Tables 3.2 and 3.3 below.

Alternative approach 1: Select a firm and build the technology landscape around it, based on the products the firm sells and the direct investments it makes. Table 3.2 lists the two databases that contain the relevant information.

Table 3.2: Financial Data Sources

Source	Data description
Capital IQ	Financial data
OECD	Country Statistics

Alternative approach 2: Select a technology domain and build a patent network. The patent network can be based on citations or on assignee, as summarized in

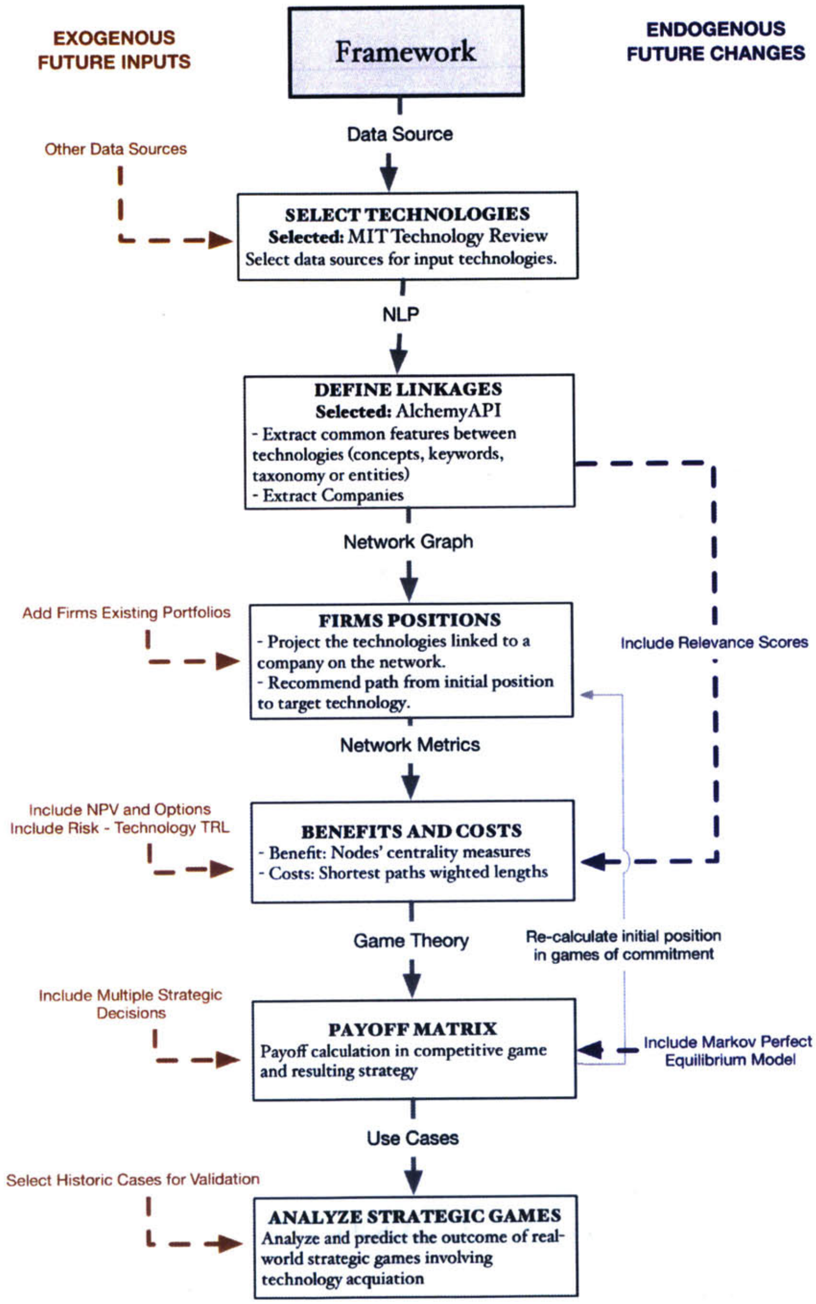

Figure 3-1: Framework Steps

Table 3.3.

Table 3.3: Patent Data Sources

Source	Data description	Technologies
PatSnap	Patent based on domains	sub-set of USPOT
World patent database	Global patent by assignee or domain	USPOT

Step 2: Define linkages between the nodes based on similarities

Once the list of breakthrough technologies is compiled, we need to define the dimensions along which to connect these technologies. Such dimensions can include: (1) patent citations; (2) functions; (3) concepts; (4) application domains; (5) common parts and components, or (6) keywords. Any one of these dimensions can generate a different network, representative of the connectivity of the technologies along a given dimension. The advantages and disadvantages of selecting these dimensions are discussed in the section (b) below.

According to Barbasi [21], it is important to choose carefully the entities represented in a network, to be able to successfully apply network theory to solve a given problem. In particular, the choice of what the links represent is a key decision in this case as there are no physical or direct connections between the technologies selected.

Links between components are more obvious when dealing with components inside a system, or with defined interactions between groups of people. For example, the technology infusion framework proposed by [11] defined connections between products or systems along four dimensions: (1) physical connection; (2) mass flow; (3) energy flow, and (4) information flow.

For the purpose of this research, the links between technologies represent similarities. Two technologies directly connected to each other would indicate shared functionality, application domain or underlying science and knowledge. Technologies farther apart would indicate differences along these dimensions.

Different dimensions that can be used as proxy to technology similarity are described below in Table 3.4.

Table 3.4: Possible Dimensions of Similarity

Dimension	Source	Description	Advantages	Disadvantages
Patent citations	List of patents (per domain or assignee)	List of patents, parents and children or assignee	Comprehensive, traceable	Sub-elements are patentable but in most cases breakthrough technologies are not represented by patents
Functions	Relies on an architectural decomposition of the technology	Proposing a functional decomposition of the technologies based on a 5X3 matrix of processes and operands	Fundamental connections of technologies based on their functions	Non-scalable and non-systematic
Concepts	NLP extraction	Using an IBM Watson suite API, extract the concepts behind the technology	Concepts represent a cross-section of the socio-technical model of the technology, finding links beyond the technical functions	NLP relies on databases of commonly defined concepts and may not be representative of novel scientific applications

Dimension	Source	Description	Advantages	Disadvantages
Application domains	These represent traditional technology domains	Selecting one domain, examining journal publications and extracting related technologies	It captures a broader scope in each domain	It is not representative of cross-domains technologies
Parts and sub-components	Bills of materials	Using a representation of the form of the technology, extract the components and parts	It can link technologies based on procurement of sub-parts	The list of components for a given technology is not always available especially when it is a process
Keywords	NLP extraction	Using an IBM Watson suite API, extract the keywords from the semantic description of the technology	Keywords are a good representation of the socio-technical environment in which the technology is used	NLP relies on databases of commonly defined concepts and may not be representative of novel scientific applications, also the list of keywords could be biased by the text description chosen

In Table 3.5 the sources and dimensions are summarized, so we can regard these options as a matrix of sources of technologies and dimensions for connectivity.

Table 3.5: Publication Sources versus Technology Dimensions

Dimensions/Sources	MIT Technology Review	World Economic Forum	OECD
Patent citation	-	-	-
Functions	X	-	-
Concepts	X	X	-
Application domains	-	X	X
Parts and sub-components	-	-	-
Keywords	X	X	-

Step 3: Represent firms' positions on the network

Once the network of breakthrough technologies is built (nodes selected and representative dimensions for linking set), the next step is to characterize a firm's position and options. For a given firm, the starting point is to identify its technology portfolio, represented by nodes in the network. The first step of the analysis is a qualitative characterization of the positions of the nodes in the network, in terms of clustering in one area of the network versus distribution and distance between nodes.

It is important to note that the portfolio of the firm does not necessarily represent the technologies owned by the firm, but rather technologies the firm already has access to through its ecosystem. The nodes in the portfolio are called the source nodes.

Step 4: Evaluate benefits and costs of nodes and edges on the network

To be able to use the network for strategic decisions we need to extract quantitative measures that characterize the nodes and links.

In particular, nodes' centrality measures are important as they represent technologies that are focal points in the network. This measure of importance is represented by three distinct node measures; degree, closeness and betweeness. These metrics represent the benefit a node holds and help identify the nodes that are likely to be sought. In Table 3.6 below, we define the network metric from a mathematical point of view, and also from its strategic implications.

Table 3.6: Networks Metrics' Strategic Interpretation

Network metric	Definition	Strategic implication
Degree	The number of connected nodes	A technology connected to many other technologies allows the firm to expand into other technology domains. Higher degrees represent options to invest in other technologies and decrease risk. Inversely, if a node is not connected to any other node, it represents a higher risk since it cannot be diversified. Higher degree value is better

Network metric	Definition	Strategic implication
Closeness	It is the inverse of farness, calculated as the sum of the distances to other nodes, and represents how quickly other nodes can be reached from a given node.	Closeness could represent a measure of cost to reach a node far from the source node. Closeness represents a measure of cost since the further away a target node is, the higher the cost. Higher closeness is better.
Betweeness	According to Ref "it is the extent to which a node is part of transactions among other nodes", in other words, it represents technologies that are part of a transition between a source and target node.	The strategic implication of this measure is that it represents enabling or generic technologies needed to move from one technology to another. In this case as well, higher values are better.
Edges weight	Represents how strongly connected two nodes are. In this case it is proportional to the number of concepts or keywords in common between the two technologies.	Strategically speaking, higher weight represents easier transition from one node to another. Cost can be taken as inversely proportional to the weight of the edge.

The benefit of a node i is the sum of degree, closeness and betweeness [33];

$$Benefit_i = Degree_i + Closeness_i + Betweeness_i \tag{3.1}$$

For the links, the path length and in particular the shortest path length, represent the cost to move between two nodes. For each pair of nodes (source and target) the shortest path length may correspond to multiple actual paths in the network (with different intermediate nodes). If the shortest path is not unique, a recommended shortest path is provided that includes the most visited intermediate nodes. The unweighted shortest path length is calculated by adding the number of links between the source node and target node and each link represents 1.

In this case we are considering a weighted network where the link between two nodes represents the number of items in common along a given dimension. More specifically, the dimension can be; concept, keyword, taxonomy or entity-company. For example, in the case of concepts the link strength corresponds to the number of individual concepts (more generally, items) that the two technologies have in common. Finally, the weight is the inverse of the strength as the shortest path length algorithm ([34] used selects the links with the lowest weight number.

Example: If nodes i and j are linked by a single segment and share 3 concepts in common, the the weight of the path is

$$w_i = \frac{1}{A_{ij}} = \frac{1}{A_{ji}} \tag{3.2}$$

$$A = \begin{array}{cc} & \begin{array}{ccc} i & j & k \end{array} \\ \begin{array}{c} i \\ j \\ k \end{array} & \begin{vmatrix} 0 & 3 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} \end{array}$$

Step 5: Evaluate the payoffs of different strategic games in the network

The network graph and network measures can now be used to quantitatively evaluate the payoff of different strategies the firm can take. For a given firm, a strategy is defined as a decision to move from an initial position to another position on the network. The initial position can be any one of the nodes in the firm's portfolio and is called the source node; the new position can only be one that is not part of the firm's portfolio and is called target node. Each pair of source-target nodes, and the specific path taken to move between the two, is a distinct strategy. The path is defined by the shortest path if it is unique, otherwise it is defined by the recommended shortest path from Step 4.

The payoff of a strategy s for company m is the difference between the benefit of the target node j and cost incurred to reach it (represented by the length l of the recommended path between the source node i and target node j). A strategy s(i,j,l) is a function of i, j, l

$$Payoff_m(s_{mk}, c) = \frac{Benefit_{j_n}}{c} - l_{i_m j_n} \tag{3.3}$$

with: c the number of competitors seeking the target j, excluding company n which is the current owner of technology j.

$i_m = 1_m ... I_m$, I_m is the total nodes in the portfolio of company m.

$m = 1...M$, M total number of companies in the landscape.

$s_{mk} = s_{m1}...S_{mK}$, K total number or strategies for company m.

$n \neq m$

Every single strategy has a different payoff. In case of competition between two companies m=1 and m=2, the payoff is modified and the benefit is reduced. In the simple case where the number of competitors is known, for example, two, the benefit is divided by 2.

In cases of uncertainty regarding the number of direct competitors, we can assign factor c in the payoff equation an estimated number of likely competitors. This can be done in various ways:

- The total number of companies n directly linked to company m (although this may simply imply these companies are in the value chain, not necessarily competitors).
- The companies n that are already present in the technology domain targeted by company m. (This means identifying the companies linked to the technologies surrounding the target technology j).
- The companies m for which the target node j_n is also highly desirable. The desirability of the node is its rank when all possible target nodes' payoffs are compared.

Continuing with company m and comparing the alternative strategies to reach the target node j_n belonging to a different company n from any node within the portfolio of m defined as i_m. The strategy s_mk recommended is the one that maximizes the payoff_m.

$$s^*_m, \quad \text{such as } Payoff_m(s^*_{mk}) = Max(Payoff_m(s_{mk})), \quad \text{for } s_{mk} = s_{m1}...s_{mK}. \tag{3.4}$$

Since the target node j_n is fixed and thereby the maximum possible benefit is fixed (although adjusted by the number of competitors involved), the variable to minimize is the cost represented by the length $l_{i_m j_n}$ between node i and j.

The method used compares all the alternative paths between

$$l^*_{i_m j_n} and \quad i_m, \quad \text{such as } l^*_{i_m j_n} = Min(l_{i_m j_n}), \quad \text{for } i_m = 1...I_m \text{ and a given } j_n. \tag{3.5}$$

The proposed implementation of the network metrics in the strategic game:

- Given a target technology j_n belonging to company n is different from m
- Given a company m=1
- Given a total number c of competing companies
- Find the list of technologies belonging to the company m, i_m
- Find the shortest path length between all i_m and j_n
- Calculate the benefit of the node j_n under competition by dividing by the number of competitors.

- Calculate the cost of the path between i_m and j_n for all i_m nodes in the portfolio of company m and the fixed target j_n.
- Select the source node i_m that leads to the smallest length between i_m and j_n and use this length as the cost of the path.
- Calculate the payoff for each company as (node benefit - path cost)

We can formalize this game in a payoff matrix shown in Table 3.7

Table 3.7: Generalized Payoff Matrix

Company m/ Competitor x	x Invests in new technology j_n	x Does not Invest in new technology j_n
m Invest in new technology j_n	$Payoff_m(s_m^*, c)), (Payoff_x(s_x^*, c)$	$Payoff_m(s_m^*, 1)), (0)$
m Does not Invest in new technology j_n	0 , $Payoff_x(s_x^*, 1)$	0 , 0

Step 6 : Apply to use cases

The use cases this framework can be applied to can be vastly different but we need to be able to define, at a minimum, the following information:

- At least one company is already present in the companies' landscape
- The company is linked to at least one technology in the technologies' landscape - At least one target new technology (at a time)
- For the competitive game, at least one competitor at a time

3.2 Data

As described in Step 1 above, there are multiple approaches to selecting the list of technologies to be included in the network. Considering advantages and disadvantages of the different sources of data presented below, the *MIT Technology Review* publication of breakthrough technology was selected as it provides a consistent description of technologies, it includes a socio-technical description of the technologies and covers a reasonable duration (2001-2016).

In the following section, an overview of each of the data sources considered is presented. For illustration purposes, the complete tables are shown in Appendix A (Data) and only an overview of the first few rows is shown in this section. In all tables the name, year and URL link of the technology are extracted. The URL link to the text description of the technology is important in the Natural Language Processing step, as the IBM NLP tool AlchemyAPI takes the URLs as inputs to perform semantic extraction.

Table 3.8 - Technologies extracted from the *MIT Technology Review* [35] : The columns represent the technology labeled 'Product', the year of selection labeled 'Year' and the URL link to the article describing the technology labeled 'Link'. the full list is provided in Appendix A.

Table 3.8: Technologies extracted from the MIT Technology Review

	Product	year	Link
0	Immune Engineering	2016	https://www.technologyreview.com/s/600763
1	Precise Gene Editing in Plants	2016	https://www.technologyreview.com/s/600765
2	Conversational Interfaces	2016	https://www.technologyreview.com/s/600766
3	Reusable Rockets	2016	https://www.technologyreview.com/s/600767
4	Robots That Teach Each Other	2016	https://www.technologyreview.com/s/600768
5	DNA App Store	2016	https://www.technologyreview.com/s/600769
6	SolarCity Gigafactory	2016	https://www.technologyreview.com/s/600770
7	Slack	2016	https://www.technologyreview.com/s/600771
8	Tesla Autopilot	2016	https://www.technologyreview.com/s/600772
9	Power from the Air	2016	https://www.technologyreview.com/s/600773

Table 3.9 - Technologies extracted from Wikipedia [36]: In this case the years of technology selection are not available, but the technology domain labeled 'Field' is provided. The

number of technologies extracted from this source is 273 and for many technologies there is an overlap with the list in Table 7.

Table 3.9: Technologies extracted from Wikipedia

Product	Link	Field
Agricultural robot	https://en.wikipedia.org/wiki/Agricultural_robot	Agriculture
Closed ecological systems	https://en.wikipedia.org/wiki/Closed_ecologica...	Agriculture
In vitro meat	https://en.wikipedia.org/wiki/In_vitro_meat	Agriculture
Precision agriculture	https://en.wikipedia.org/wiki/Precision_agricu...	Agriculture
Vertical farming	https://en.wikipedia.org/wiki/Vertical_farming	Agriculture
Drones	https://en.wikipedia.org/wiki/Unmanned_aerial_...	Aviation
Micro air vehicle	https://en.wikipedia.org/wiki/Micro_air_vehicle	Aviation
Neural-sensing headset	https://en.wikipedia.org/wiki/Honeywell	Aviation
Atmospheric carbon dioxide removal	https://en.wikipedia.org/wiki/Carbon_dioxide_r...	Climate engineering
3D printing	https://en.wikipedia.org/wiki/3D_printing	Construction

Table 3.10 - Technologies extracted from the World Economic Forum's reports (2014-2015) [37] and [38]: This list is compiled from the World Economic Forum's publication of breakthrough technologies for 2014 and 2015. It contains ten technologies per year and instead of providing the url, the text description is provided. This is the other input type that AlchemyAPI takes for semantic extraction.

Table 3.10: Technologies extracted from the WEF

Technology	Description	Year
Body-adapted Wearable Electronics	From Google Glass to the Fitbit wristband, wea...	2014
Screenless Display	One of the more frustrating aspects of modern ...	2014
Human Microbiome Therapeutics	The human body is perhaps more properly descri...	2014
RNA-based Therapeutics	RNA is an essential molecule in cellular biolo...	2014
Quantified Self (Predictive Analytics)	The quantified-self movement has existed for m...	2014
Brain computer Interfaces	The ability to control a computer using only t...	2014
Nanostructured Carbon Composites	Emissions from the world s rapidly-growing fle...	2014
Mining Metals from Desalination Brine	As the global population continues to grow and...	2014
Grid-scale Electricity Storage	Electricity cannot be directly stored, so elec...	2014
Nanowire Lithium-ion Batteries	As stores of electrical charge, batteries are ...	2014

Breakthrough technologies and technology domains Technology domains relate to the different domains of expertise. The official definition of these domains is taken from the National Bureau of Economic Research [39]

Classification of technology in a given domain can be subjective and at the same time it is highly important to assess technology connections across-domains. Classification of technologies are mostly based on patents as they are readily available and already include a classification. The intent of this work is to propose a classification based on technology similarity.

Technology opportunity discovery studies have led the way in the area of technology similarity analysis. Some studies focused on promising new technologies in given areas [40] and others tried to tie these new technologies to an existing portfolio of technology and products at a firm level [5]. In this paper, Yoon explores a NLP approach for Subject-action-object (SAO) extraction that can derive functional similarities. The study further proposes a framework capable of identifying technology opportunities consistent with an existing technology portfolio of a company.

While this method removes the subjectivity challenge of assigning a technology to a technology domain, it relies on the existence of patents related to the technology. This in itself means the technology needs to be first decomposed into sub-elements, which relies on domain expertise.

A special case of semantic extraction (Subject - Action - Object) SAO: IBM's AlchemyAPI NLP tool is able to extract SAO structures from the parsed text. The preliminary review of the SAO elements revealed that they represent longer sentence structures as shown in the example below. While in principle longer structures can be further decomposed through iterations to isolate the SAO elements of interest. And from that derive the functional decomposition of a technology. This method is better suited if the text description is focused on a technical description of the technology rather than a general article that includes socio-technical descriptions.

Table 3.11: SAO Examples

	Type	Relations
0	Action	['was', 'was announced', 'was', 'runs', 'had',...
0	Object	['announced', 'unimpressed', 'a payment startu...
0	Subject	['Apple Pay', 'Apple Pay', 'Osama Bedier', 'A ...
1	Action	['lifts', 'bob', 'are', 'possess', 'are', 'cal...
1	Object	['a clear plastic dish', 'cerebral organoids, ...
1	Subject	['Madeline Lancaster', 'tissue the size of sma...
2	Action	['looks', 'reminds', 'was', 'was speeding', 'w...
2	Object	['like a street racer', 'of a math teacher', '...
2	Subject	['Hariharan Krishnan', 'he', 'he', 'he', 'I', ...

Chapter 4

Technology characterization

4.1 Technology Classification

Technology classification is important to understand relationships between technologies. The difficulty is consistency across disciplines. The most largely used type of classification is the one related to patents categories provided by the USPTO. This classification schema consists of 3-digit patent classes with 120,000 sub-classes. The National bureau of economic research [39] proposed an aggregation into 36 two-digit technological sub-categories themselves aggregated into 6 main categories (Chemical (excluding drugs); Computers and Communications (C&C); Drugs and Medical (D&M); Electrical and Electronics (E&E); Mechanical; and others. According to the authors of the study, this classification still exhibits inherent "arbitrariness" and "should be examined critically for specific applications".

(Include in Appendix the list from NBER Appendix A)

By comparison, other technology classification domains are based on high level functional descriptions . Such a classification is presented in Reference [14]. The 5 by 3 matrix decomposes technology domains based on processes and operands as shown in Figure 4-1 below.

Table 1
Functional technological classification with operands and operations [17,18]

Operation	Operand		
	Matter (M)	Energy (E)	Information (I)
Transform	Blast furnace	Lamps, electrical generator	*Analytic engine, Calculator*
Transport	Truck	Electrical grid	*Cables, Radio, Television*
Storage	Warehouse	Batteries, flywheels	*Magnetic tape and disk, Book*
Exchange	eBay Trading System	Energy markets	World wide web, wikipedia
Control	Health Care System	Atomic energy commission	Internet engineering task force

Figure 4-1: Functional Technology Classification

Later, (Do, 2014) [15] proposed an extension of the classification table to include "Living Matter" and "Money" as elementary operands and resulted in a 5 by 5 matrix.

The classification of any given technology in one of these categories is also subject to variations and subjectivity. Starting with the data source of technologies selected in Chapter 2, we use a simple classification algorithm on a small training data set and run it on the remaining technologies to classify them. In this case we are using a python library of machine learning algorithms built for processing textual data [41]. In particular the classifier used is based on a naive Bayes algorithm [42].

For instance staring with a training set of 16 technologies representing about 10% of the total data set, we assign the appropriate classification to both operands and processes. This is shown in Table tab:tab5-1.

Table 4.1: Example 5 by 5 matrix of Processes and Operands

Process Operand	Living Matter	Matter	Energy	Information	Matter
Transform	-	-	-	-	-
Transport	-	-	-	-	-
Store	-	-	-	-	-
Control	-	-	-	-	-
Exchange	-	-	-	-	-

The training set selected is shown on Table 5.2. We note that a number of (process,operand) pairs are not represented in the table as there are no corresponding technology breakthroughs in this category.

Table 4.2: Processes and Operands Classification Training Set

Process Operand	Living Matter	Matter	Energy	Information	Money
Transform	“Immune Engineering”	“microscale 3-D printing”	smart transformers	“Robots That Teach Each Other”	-
Transport	-	“Reusable Rockets”	-	“car to car communication”	“apple pay”
Store	“brain organoids”	-	“SolarCity Gigafactory”	“universal memory”	-
Exchange	-	-	“Power from the Air”	“Conversational Interfaces”, “Slack”	“crowdfunding”

Process Operand	Living Matter	Matter	Energy	Information	Money
Control	"neuron control"	-	"power grid control"	"quantum cryptography"	-

With the classification algorithm trained on the "name" of the technology, the resulting classification of the entire set of technologies is shown below in Table 4.3.

Table 4.3: Technology Classification based on Names

	Living Matter	Matter	Energy	Information	Money
Transform	[Immune Engineering]	[microscale 3-D printing]	[smart transform-ers]	[Robots That Teach Each Other, smart wind and ...	NaN
Transport	NaN	NaN	NaN	[Reusable Rockets, car to car communication]	[apple pay]
Store	[brain organoids]	NaN	[SolarCity Gigafactory]	[brain mapping, memory implants, racetrack mem...	NaN
Exchange	NaN	NaN	[Power from the Air]	[Precise Gene Editing in Plants, Conversational...	NaN
Control	[neuron control]	NaN	[power grid control]	[quantum cryptogrpahy]	NaN

Recognizing the limits of the training and classification solely based on the name of the technology we have also tested a method where the classifier is trained on the concepts linked to the technology and is able to classify a new technology based on its concepts. We can see an improvement on the classification presented in Table 4.4 below as compared to Table 4.3. In particular we can see that in the category "Energy" and "Transport" we are able to capture both "Reusable Rockets" and "Tesla Autopilot".

Table 4.4: Technology Classification based on Concepts

	Living Matter	Matter	Energy	Information	Money
Transform	[Immune Engineering, DNA App Store, internet d...	[microscale 3-D printing, synthetic cells, sin...	[genome editing, high speed material discovery...	[Precise Gene Editing in Plants, Robots That T...	NaN
Transport	NaN	[Reusable Rockets, Tesla Autopilot]	[airborne networks, mechatronics]	[car to car communication, nano-architecture, ...	[apple pay]
Store	[brain organoids, brain mapping, deep learning...	NaN	[SolarCity Gigafactory, super grids, ultra-effi...	[hashCache, racetrack memory, offline web appl...	NaN
Exchange	NaN	NaN	[Power from the Air, solar micro-grids, nanora...	[Conversational Interfaces, Slack, magic leap,...	[crowdfunding]
Control	[neuromorphic chips, biological machines, conn...	NaN	[agricultural drones, oculus rift, smart wind ...	[3-D transistors, homomorphic encryption, envi...	NaN

Next we evaluate the classification using extracted keywords rather than concepts, the results are presented in Table 4.5 below. The machine learning algorithm is trained on the individual keywords linked to a given technology and the classification happens on the keywords associated with the new technology.

Table 4.5: Technology Classification based on Keywords

	Living Matter	Matter	Energy	Information	Money
Transform	[Immune Engineering, liquid biopsy, microscale...	[nuclear reprogramming, microfluidics]	[supercharged photosynthesis, synthetic cells,...	[Precise Gene Editing in Plants, Robots That T...	NaN
Transport	NaN	[Reusable Rockets, Tesla Autopilot]	[biological machines, grid computing, softwar...	[apple pay, car to car communication, agricult...	NaN
Store	[brain organoids, neuromorphic chips, memory i...	NaN	[SolarCity Gigafactory, smart wind and solar, ...	[facebook timeline, racetrack memory, graphene...	NaN
Exchange	NaN	NaN	[Power from the Air, mobile 3-D, pervasive wir...	[Conversational Interfaces, Slack, magic leap,...	[crowdfunding, cell-phone viruses]
Control	[brain mapping, connectomics, personalized med...	[DNA App Store]	[agile robots, homomorphic encryption, traveli...	[cloud programming, real-time search, nanowire...	NaN

In this section we used the technology names, concepts and keywords extracted from the the data source of 150 technologies to classify these technologies in a 5 by 5 matrix representing operands and processes. In the next section we will use these same extracted features (concepts and keywords) as well as taxonomy and entities to analyze technologies and companies inter-connectedness.

4.2 Technologies and Companies Inter-Relatedness

The list of technologies used here is extracted from the data source discussed in Chapter 3. Starting from the 150 technologies we extract their concepts, keywords, taxonomy and entities. In this case we limit entities to the companies linked directly to the technologies. Table 4.6 below shows the list of networks built and analyzed in this section and some of their characteristics. The objective of this section is to identify the most relevant dimensions of linkage among technologies and select the best representation of the technology network that can be used in Chapter 5 for the simulating strategic games.

Table 4.6: Characteristics of the Networks Analyzed

Main Networks	Nodes	Links	Directed	Size
Technology linkage through concepts	Technologies	Concepts	undirected	149 nodes, 971 links
Technology linkage through keywords	Technologies	Keywords	undirected	150 nodes, 533 links
Technology linkage through taxonomy	Technologies	Taxonomy	undirected	150 nodes, 8417 links
Technology linkage through entities(companies)	Technologies	Entities	undirected	150 nodes, 2584 links
Companies linkage through concepts	Companies	Concepts	undirected	229 nodes, 4067 links
Companies linkage through entities(companies)	Companies	Entities	undirected	229 nodes, 525 links

4.2.1 Technology to Technology Networks

Technology linkage through concepts represent an abstraction of the socio-technical descriptions of the technologies. In this section we will analyze four networks with the same number of nodes representing the 150 technologies selected and varying numbers of links depending on the dimension analyzed. We then compare their metrics and select the most appropriate representation.

For each one of the four networks we calculate the degree, betweeness and closeness of each node as described in Chapter 1. The total score assigned to the node is the sum of its degree, betweeness and closeness. While the the betweeness and closeness are normalized the degree is not and it therefore dominates the score value of a node. This approach gives more weight to the degree of a node which represents a technology diverse applicability potential as it connects to many other technologies.

Betweeness is an indication of whether the node (technology) is an intermediate step in paths connecting a pair of nodes. Higher values indicates that there many shortest paths that go through this node to link and inversely lower values are an indication that the node does not intervene in many shortest paths. Betweeness can be interpreted as the capability of a node to block a path, thereby preventing a node from becoming reachable or forcing the path to be longer (non-optimum).

Closeness is an indication of how far away other nodes are from the given node. The higher the closeness the shortest it is to reach the other nodes. The desirability of such nodes is high and as we will see in the following sections the correlation of closeness to degree is high.

In the following sections we will first explore in details the case of the network built on concepts. Then the following three networks are the summarized and compared to the concepts network.

1. Technology linkage through Concepts The network of technologies analyzed in this section include 149 nodes, the technology "Modeling Surprise" was excluded from the sample as the description indicated it is an approach rather than a process or technology. The following important network metrics are calculated and summarized in Table 4.7.

Table 4.7: Network Metrics for the Concept-based Technology Network

Measure	Value
Number of nodes	149
Number of edges	971
Average degree	13.03
Number of connected components	7
Size of largest connected component	143
The average shortest path length	2.79

In addition to the summary table the we build a centrality table that summarizes for each node the degree, betweeness, closeness and total score. The first five rows are shown below in Table 4.8 .

Table 4.8: Nodes' Centrality Measures for the Concept-based Technology Network

node	degree	betweenness	closeness	score
Precise Gene Editing in Plants	25	0.004955	0.377660	25.382615
graphene transistors	14	0.018532	0.449367	14.467899
racetrack memory	16	0.031878	0.426426	16.458304
grid computing	10	0.001977	0.346341	10.348319
bacterial factories	25	0.005876	0.377660	25.383536

The degree distribution histogram is shown on Figure 4-2 with both the in-degree and out-degree repre-

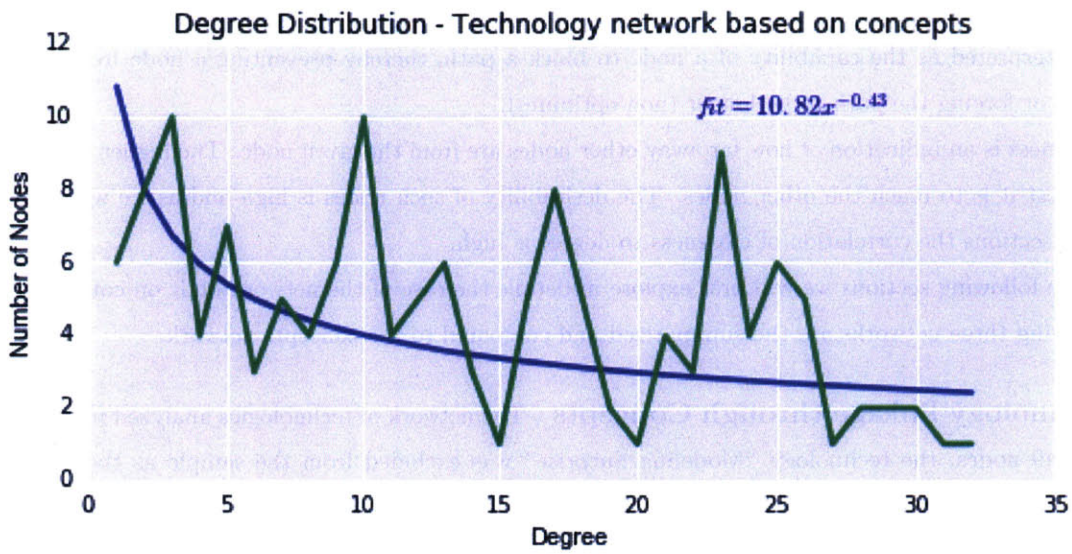

Figure 4-2: Nodes' Degree Distribution for the Concept-based Technology Network

sented even though they are the same for an undirected network. Note that the histogram does not exhibit a power law distribution. As discussed in Chapter 2, real-world networks can deviate substantially from the power law behavior seen in random networks.

We note that the correlation between degree and closeness as well as between degree and betweeness are both positive and closeness has a high correlation. This confirms the hypothesis that the node's score is well represented by the degree. Table 4.9 summarizes the correlation results of the different network measures.

Table 4.9: Centrality Measures' Correlation for the Concept-based Technology Network

	degree	betweenness	closeness	score
degree	1.000000	0.538136	0.713626	0.999987
betweenness	0.538136	1.000000	0.607686	0.520665
closeness	0.713626	0.607686	1.000000	0.717085
score	0.999987	0.520665	0.717085	1.000000

Finally, while direct links between the nodes provide information at an individual node level, we can see that a clusterization (also defined as partitioning) of the network uncovers a number of communities. In this case, five clusters are defined and are listed in Table 4.10 below.

Communities represent sets of nodes exhibiting strong inter-connectedness [43], Figure 4-3 shows the communities on the technology network.

Table 4.10: Clusterization of the Concept-based Technology Network

Cluster	Technologies
Cluster1	'Precise Gene Editing in Plants', 'supercharged photosynthesis', 'bacterial factories', 'dual-action antibodies', 'genome editing', 'DNA App Store', 'microscale 3-D printing', 'liquid biopsy', 'separating chromosomes', 'metabolomics', 'cellulolytic Enzymes', 'nanopore sequencing', 'microfluidics', 'nanimprint lithogrpahy', 'molecular imaging', 'epigenetics', 'comparative interactomics', 'nanobiomechanics', 'Immune Engineering', 'nanohealing', 'rNAi interference', 'prenatal DNA sequencing', '$100 genome', 'nanomedicine', 'internet dna', 'magnetic-resonance force microscopy', 'synthetic cells', 'cancer genomics', 'glycomics', 'nuclear reprogramming', 'personal genomics', 'single-cell analysis'
Cluster2	'graphene transistors', 'racetrack memory', 'nanocharging solar', 'solar fuel', 't-rays', 'nanowires', 'paper diagnostics', 'nano-architecture', 'green concrete', 'liquid battery', 'ultra-efficient solar power 1', 'wireless power', 'smart transformers', 'megascale desalination', '3-D transistors', 'microfluidic optical fibers', 'SolarCity Gigafactory', 'nanopiezoelectronics', 'nanoradio', 'solid state batteris', 'nanosolar cells', 'agricultural drones', 'quantum cryptogrpahy', 'smart wind and solar', 'neuromorphic chips', 'ultra-efficient solar power 2', 'flexible transistors', 'microphotonics', 'digital imaging reimagined', 'quantum wires', 'supergrids', 'high speed material discovery', 'silicon photonics', 'invisible revolution', 'a new focus for light', 'light-trapping photovoltaics', 'light field photography', 'implantable electronics', 'power grid control', 'atomic magnetometers', 'stretchable silicon', 'mechatronics'
Cluster3	'biometrics', 'reality mining', 'apple pay', 'Slack', 'intelligent software assistant', 'cognitive radio', 'Power from the Air', 'gestural interfaces', 'augmented reality', 'smart watches', 'car to car communication', 'probabilisitic chips', 'pervasive wireless', 'solar micro-grids', 'digital rights management', 'mobile 3-D', 'software defined networking', 'big data from cheap phones', 'mobile collaboration', 'Tesla Autopilot', 'ultraprivate smartphones', 'Reusable Rockets', 'oculus rift', 'social tv', 'wireless sensor networks'

Cluster	Technologies
Cluster4	'untangling code', 'personalized medical monitors', 'social indexing', 'cloud streaming', 'hashCache', 'grid computing', 'cloud programming', 'offline web applications', 'facebook timeline', 'baxter the blue collar robot', 'distributed storage', 'homomorphic encryption', 'robot design', 'Robots That Teach Each Other', 'brain machine interface', 'Conversational Interfaces', 'data mining', 'deep learning', 'cell-phone viruses', 'real-time search', "peering into video's future", 'temporary social media', 'universal translation', 'universal authentication', 'synthetic biology', 'bayesian machine learning', 'natural language processing', 'universal memory', 'sooftware assurance', 'agile robots', 'crash-proof code', 'project Loon', 'magic leap'
Cluster5	'injectible tissue engineering', 'neuron control', 'brain mapping', 'memory implants', 'biomechanics', 'egg stem cells', 'brain organoids', 'connectomics', 'engineered stem cells', 'biological machines', 'diffusion tensor imaging'

In fact one interesting view of this network is its evolution over time as shown in Figure 4-4. Looking at these same communities on a year by year basis, we are able to see the additional nodes appearing an filling the network.

We can see for instance that the "Information" cluster and the main "Living Matter" cluster evolved separately between 2001 and 2007. Furthermore, it seems that the in this initial phase the nodes are spread-out and farther and more on the edges of what will later be the complete network. During this period, the communities became denser (more nodes) and had more links within the community. In the next period, between 2008 and 2013 the communities continued to grow in density but internally within the boundaries set in the previous period. Finally, between 2014 and 2016 the links within communities started to form as well as the addition of a few new nodes on the edges.

To better understand the underlying connections between technologies, we will explore a few examples in greater details.

Example 1: the path between two technologies 'Tesla Autopilot' and 'Reusable Rockets'

The concepts extracted for 'Tesla Autopilot' and 'Reusable Rockets' are shown in Table 4.11 and 4.12 respectively. The relevance scores of the each concept extracted from the NLP analysis are also shown. The top concepts show a meaningful categorization of the two technologies in automobile and space rockets. However, we note that these concepts also include references to key people, locations and field terminology thereby capturing the technology's ecosystem. In this case there is only one concept in common, representing an important stakeholder "Elon Musk" in the two companies commercializing these technologies.

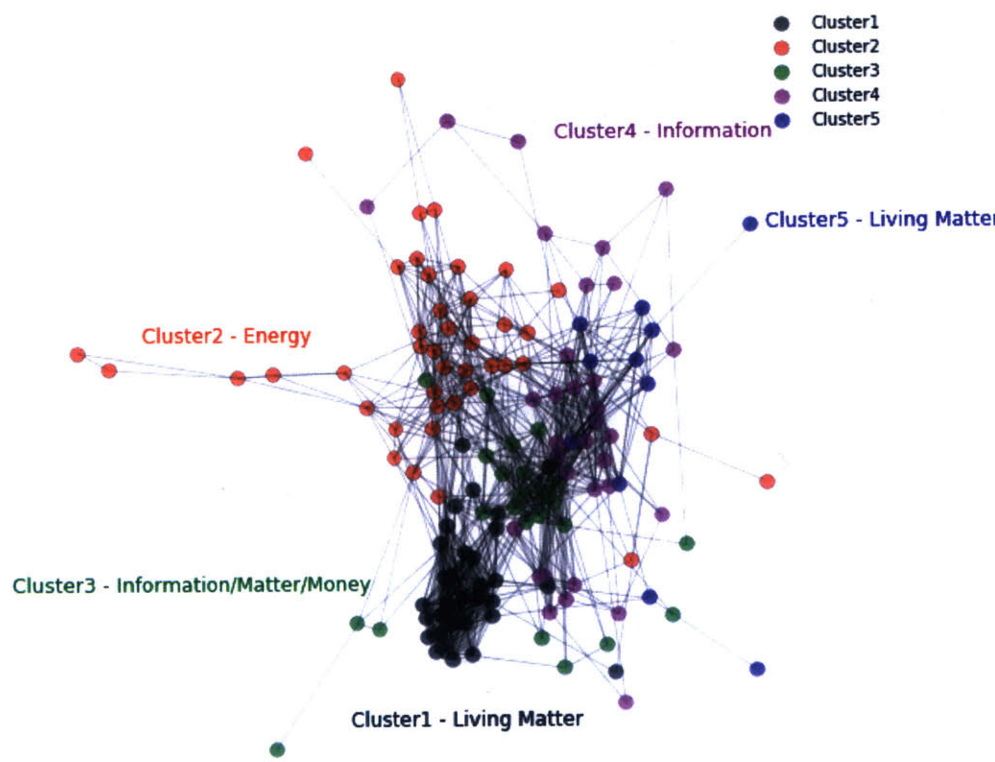

Figure 4-3: Clusterization for the Concept-based Technology Network

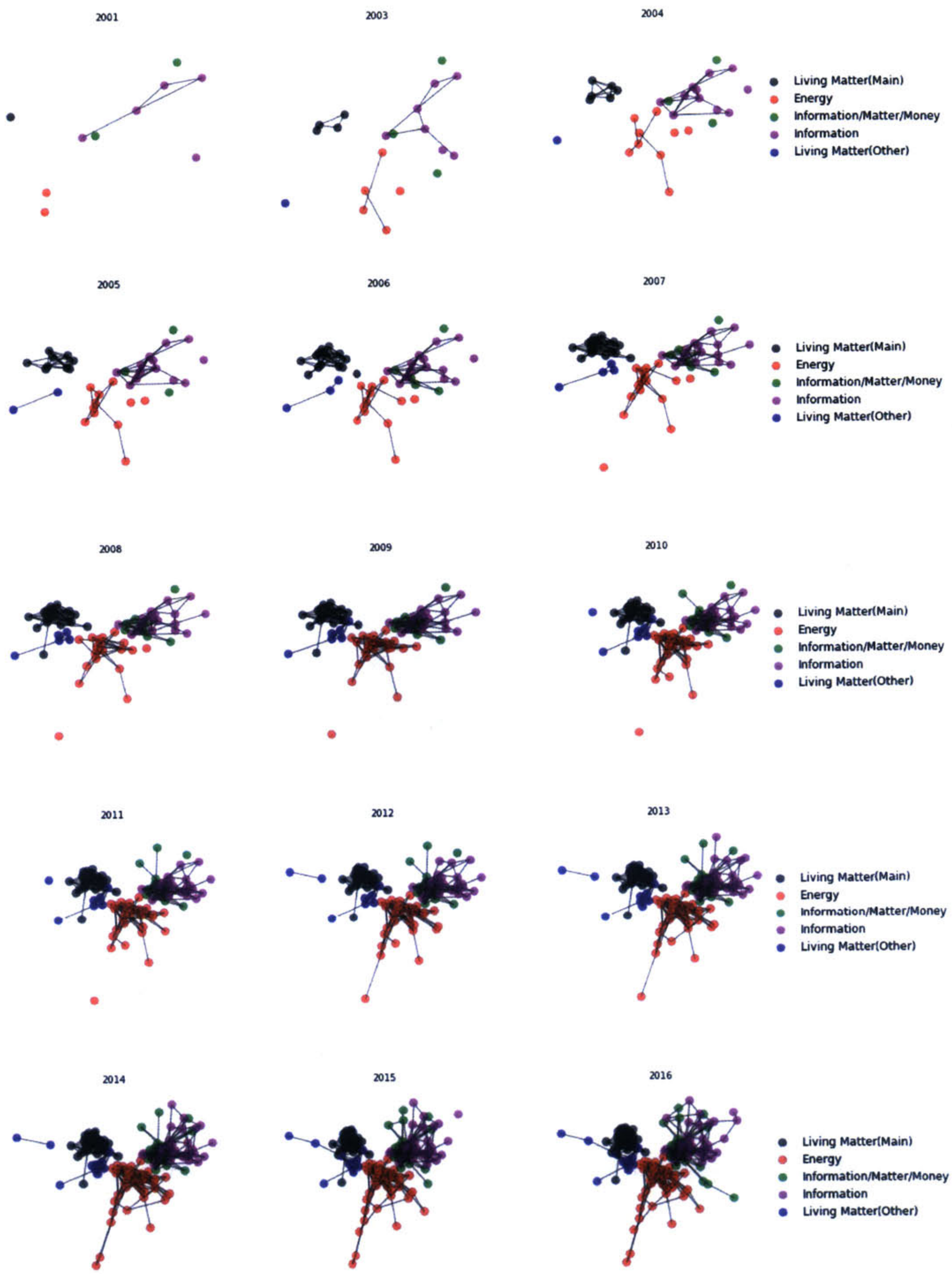

Figure 4-4: Evolution over time of the technology communities in the Concept-based Technology Network

Table 4.11: Concepts Related to 'Tesla Autopilot'

Concept	Relevance	Technology	Year
Automobile	0.915264	Tesla Autopilot	2016
Steering	0.856171	Tesla Autopilot	2016
Elon Musk	0.819283	Tesla Autopilot	2016
Opel	0.727092	Tesla Autopilot	2016
Renault	0.693996	Tesla Autopilot	2016
Chevrolet	0.673986	Tesla Autopilot	2016
Japan	0.623282	Tesla Autopilot	2016
Ultrasound	0.612602	Tesla Autopilot	2016

Concepts for 'Reusable Rockets':

Table 4.12: Concepts Related to 'Reusable Rockets'

Concept	Relevance	Technology	Year
Rocket	0.951593	Reusable Rockets	2016
Jeff Bezos	0.784745	Reusable Rockets	2016
Elon Musk	0.784277	Reusable Rockets	2016
Outer space	0.754564	Reusable Rockets	2016
Time	0.705892	Reusable Rockets	2016
Spacecraft	0.657821	Reusable Rockets	2016
Space exploration	0.58404	Reusable Rockets	2016
Stephen Baxter	0.581095	Reusable Rockets	2016

What we want to explore next are the shortest path lengths between these two technologies, using a custom python function we test the following features of this path: - existence - shortest path length - default shortest path returned by the python library used - number of links in the shortest path - all alternative shortest paths with the same length - if the shortest path is not unique, returns a recommended shortest path that includes the most visited nodes

As an example the returned output for the two technologies explored is shown below with the weighted paths shown in Table 4.13 and the shortest path highlighted on the network shown in Figure 4-5.

- There is a path from source to target? True
- Shortest path lengths in the graph: 1.0
- Automatically generated shortest path in the graph: ['Tesla Autopilot', 'Reusable Rockets']
- Number of links in the shortest path: 1
- Possible path: ['Tesla Autopilot', 'Reusable Rockets']
- Number of distinct shortest paths in the graph: 1
- Recommended Path(s): ['Tesla Autopilot', 'Reusable Rockets']

Table 4.13: Links' Weights between Tesla Autopilot' and 'Reusable Rockets' Concepts Network

Source	Target	Weighted Length
Tesla Autopilot	Reusable Rockets	1

Example 2: In this second example we look at two technologies in the bio technologies domain 'Precise Gene Editing in Plants' and 'genome editing'. Tables 4.14 and 4.15 summarize the concepts and their relevance. Note that in this case there are 3 concepts in common; DNA, Gene and Genetics. The higher number of concepts in common is an indication of greater similarity between the technologies.

Concepts for 'Precise Gene Editing in Plants':

Table 4.14: Concepts Related to 'Precise Gene Editing'

Concept	Relevance	Technology	Year
DNA	0.949147	Precise Gene Editing in Plants	2016
Bacteria	0.833692	Precise Gene Editing in Plants	2016
Genetically modified food	0.731114	Precise Gene Editing in Plants	2016
Gene	0.727197	Precise Gene Editing in Plants	2016
Organism	0.679589	Precise Gene Editing in Plants	2016
Molecular biology	0.67392	Precise Gene Editing in Plants	2016
Seed	0.645407	Precise Gene Editing in Plants	2016

Concept	Relevance	Technology	Year
Genetics	0.639032	Precise Gene Editing in Plants	2016

Concepts for 'genome editing':

Table 4.15: Concepts Related to 'genome editing'

Concept	Relevance	Technology	Year
DNA	0.956925	genome editing	2014
Genetics	0.783562	genome editing	2014
Mutation	0.620114	genome editing	2014
Gene	0.576466	genome editing	2014
Primate	0.559777	genome editing	2014
In vitro fertilization	0.522782	genome editing	2014
Genetic engineering	0.507729	genome editing	2014
Genetic disorder	0.456013	genome editing	2014

Following the same approach as for the previous set of technologies we are able to verify the existence of a path, extract its length, the default shortest path returned by the program, the number of links in the shortest path as well as the alternative paths and recommended path. Figure 4-6 shows network with the recommended path. Table 4.16 shows the weight of the links along the shortest path. The full output of the program is shown below.

- There is a path from source to target? True
- Shortest path lengths in the graph: 0.333333333333
- Automatically generated shortest path in the graph: ['Precise Gene Editing in Plants', 'genome editing']
- Number of links in the shortest path: 1
- Possible path: ['Precise Gene Editing in Plants', 'genome editing']
- Number of distinct shortest paths in the graph: 1
- Recommended Path(s):['Precise Gene Editing in Plants', 'genome editing']

Table 4.16: Links' Weights between Precise Gene Editing in Plants' and 'genome editing' Concepts Network

Source	Target	Weighted Length
Precise Gene Editing in Plants	genome editing	0.333333

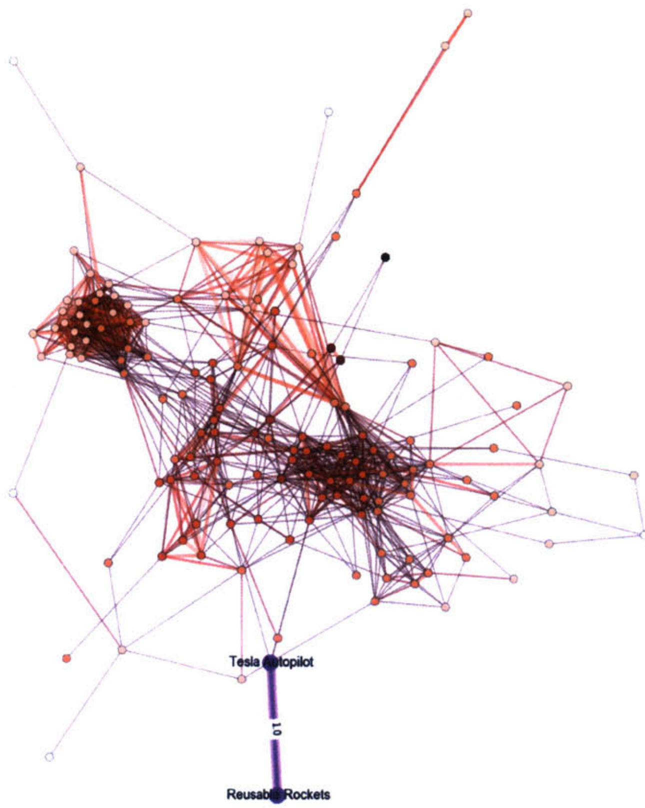

Figure 4-5: Shortest Path between 'Tesla Autopilot' and 'Reusable Rockets'

Figure 4-6: Shortest Path between 'Precise Gene Editing in Plants' and 'genome editing'

Example 3: In this example we have selected two technologies that we expect to be very different from each other 'magic leap' in Cluster 4 (Information) and 'solar fuel' in Cluster 2 (Energy).

> Magic Leap 2015: Magic leap is a virtual reality technology that uses a device to make virtual objects appear in real life. This technology is revolutionary for a number of domains especially in the film and gaming industries. The device uses a small projector that shines light onto a transparent lens. The reflected light has a pattern that blends with the surrounding light. This process tricks the visual cortex in a way that makes artificial objects indistinguishable from real objects [44].
>
> Solar Fuel 2010: Solar fuel is a technology that uses sunlight to efficiently convert carbon dioxide into ethanol or diesel. This technology is based on the principle that bio fuels can be

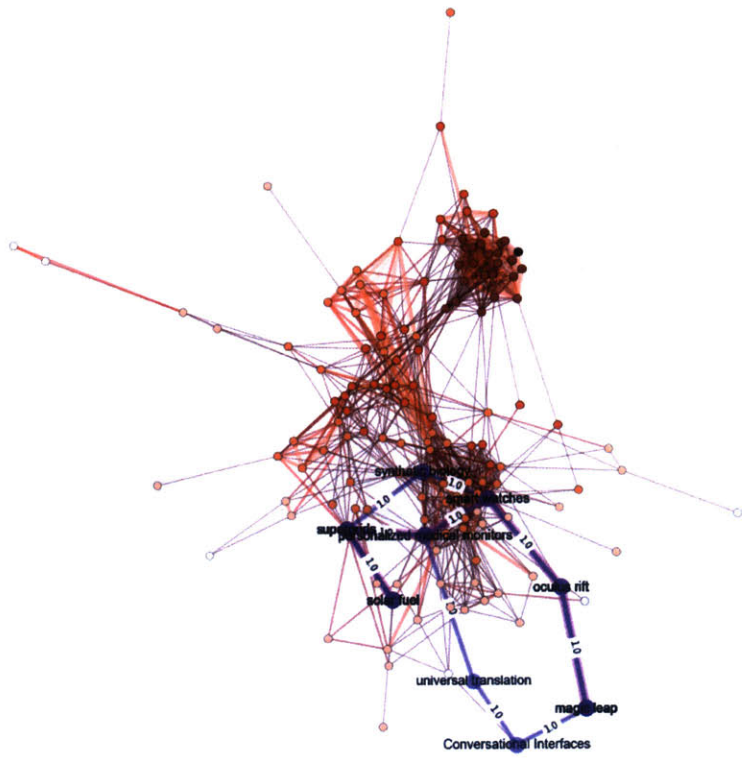

Figure 4-7: Shortest Paths between 'magic leap' and 'solar fuel' - Concepts

directly generated from carbon dioxide and water and the use of biomass such as corn or switch grass or algae can be eliminated since they are an intermediate step. Solar fuel achieves this by manipulating genes to create photosynthetic micro-organisms [45].

In this case, there are 3 possible paths between the source node 'magic leap' and the target node 'solar fuel' with the a path length of 5.0 equal to the shortest path length.

The three paths cross some of the nodes multiple times making them the most visited nodes between this source-target pair. Therefore, the recommended shortest path includes these nodes. The three alternative shortest paths are shown in blue on Figure 4-7 and the recommended path is highlighted in magenta. Tables 4.17 through 4.19 summarize the weights of the links along the alternative shortest paths.

- There is a path from source to target? True

- Shortest path lengths in the graph: 5.0
- Automatically generated shortest path in the graph: ['magic leap', 'oculus rift', 'smart watches', 'personalized - - - medical monitors', 'supergrids', 'solar fuel']
- Number of links in the shortest path: 5
- Possible path: ['magic leap', 'oculus rift', 'smart watches', 'personalized medical monitors', 'supergrids', 'solar fuel']
- Possible path: ['magic leap', 'Conversational Interfaces', 'universal translation', 'personalized medical monitors', 'supergrids', 'solar fuel']
- Possible path: ['magic leap', 'oculus rift', 'smart watches', 'synthetic biology', 'supergrids', 'solar fuel']
- Number of distinct shortest paths in the graph: 3 recommended Path(s):['magic leap', 'oculus rift', 'smart watches', 'personalized medical monitors', 'supergrids', 'solar fuel']

Table 4.17: Links' Weights between 'magic leap' and 'solar fuel'- Alternative 1 Concepts Network

Source	Target	Weighted Length
supergrids	solar fuel	1
personalized medical monitors	smart watches	1
magic leap	oculus rift	1
personalized medical monitors	supergrids	1
oculus rift	smart watches	1

Table 4.18: Links' Weights between 'magic leap' and 'solar fuel'- Alternative 2 Concepts Network

Source	Target	Weighted Length
Conversational Interfaces	universal translation	1
personalized medical monitors	universal translation	1
personalized medical monitors	supergrids	1
magic leap	Conversational Interfaces	1
supergrids	solar fuel	1

Table 4.19: Links' Weights between 'magic leap' and 'solar fuel'- Alternative 3 Concepts Network

Source	Target	Weighted Length
supergrids	solar fuel	1

Table 4.19: Links' Weights between 'magic leap' and 'solar fuel'- Alternative 3 Concepts Network

Source	Target	Weighted Length
magic leap	oculus rift	1
synthetic biology	supergrids	1
synthetic biology	smart watches	1
oculus rift	smart watches	1

In the next three sections we will compare the technology networks based on keywords, taxonomy and entities to the concepts' network. We will use the same example of technology pairs (magic leap to solar fuel) to illustrate how the dimension chosen for the links may change the shortest path.

2. Technology linkage through Keywords The network generated from keywords exhibits some similarities to the one generated from concepts and discussed in the previous section. In particular, the average shortest path length and the size of the largest connected component are similar. The main differences are; a lower number of links and lower average degree. In addition, the number of connected components is larger. Tables 4.20 and 4.21 show respectively the metrics of the network and the centrality measures of the nodes.

Table 4.20: Network Metrics for the Keywords-based Technology Network

Measure	Value
Number of nodes	150
Number of edges	533
Average degree	7.37
Number of connected components	13
Size of largest connected component	138
The average shortest path length	3.14

Table 4.21: Nodes' Centrality Measures for the Keywords-based Technology Network

node	degree	betweenness	closeness	score
ultra private smart phones	7	0.006839	0.352185	7.359024
Precise Gene Editing in Plants	6	0.003963	0.314943	6.318906
graphene transistors	16	0.061190	0.378453	16.439643

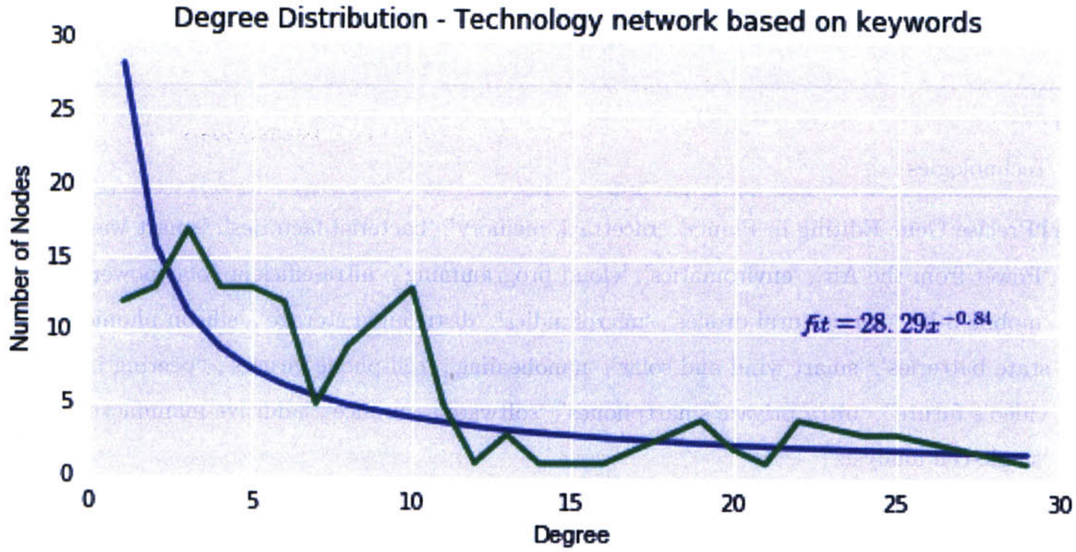

Figure 4-8: Nodes' Degree Distribution for the Keywords-based Technology Network

node	degree	betweenness	closeness	score
racetrack memory	11	0.034364	0.362434	11.396798
supercharged photosynthesis	3	0.003092	0.312073	3.315165

The degree distribution of this network is closer to a power law as shown on Figure 4-8. The power law has a parameter of 0.84 compared to 0.43 seen in the concepts' network.

Finally, when looking at the clusterization of the network, we are able to extract seven major clusters. Upon examination of the individual technologies in these clusters, we can see that the groups contain a mix of technologies that belong to different technology domains. For example: 'Precise Gene Editing in Plants' and 'cell-phone viruses' are in the same group, 'prenatal DNA sequencing' and 'graphene transistors' are in a second group.

Unlike with the network of technologies along concepts, the clusters listed here include heterogeneous technologies and the labeling or assignment of these technologies to a single technology domain is not straightforward.

The derived clusters are listed in Table 4.22 and graphically shown on Figure 4-9.

Table 4.22: Clusterization of the Keywords-based Technology Network

Cluster	Technologies
Cluster1	['Precise Gene Editing in Plants', 'racetrack memory', 'bacterial factories', 'smart watches', 'Power from the Air', 'enviromatics', 'cloud programming', 'ultra-efficient solar power 2', 'mobile 3-D', 'agricultural drones', 'microfluidics', 'distributed storage', 'silicon phonics', 'solid state batteries', 'smart wind and solar', 'nanohealing', 'cell-phone viruses', "peering into video's future", 'ultra private smartphones', 'software assurance', 'additive manufacturing', 'single-cell analysis']
Cluster2	['graphene transistors', 'dual-action antibodies', 'paper diagnostics', 'nano-architecture', 'green concrete', '3-D transistors', 'crowdfunding', 'microscale 3-D printing', 'biomechanics', 'traveling wave reactor', 'egg stem cells', 'molecular imaging', 'epigenetics', 'a faster Fourier Transform', 'smart transformers', 'microphotonics', 'quantum wires', 'prenatal DNA sequencing', 'memory implants', 'real-time search', 'injectible tissue engineering', 'universal translation', 'Tesla Autopilot', 'synthetic cells', 'implantable electronics', 'synthetic biology', 'universal memory', 'project Loon', 'magic leap']
Cluster3	['supercharged photosynthesis', 'personalized medical monitors', 'cloud streaming', 'ultra-efficient solar power 1', 'offline web applications', 'probabilisitic chips', 'homomorphic encryption', 'nanoradio', 'neuromorphic chips', 'software defined networking']
Cluster4	['untangling code', 'social indexing', 'hashCache', 'baxter the blue collar robot', 'Robots That Teach Each Other', 'nanosolar cells', 'nanobiomechanics', 'quantum cryptogrpahy', 'solar micro-grids', 'Immune Engineering', 'nanocharging solar', 'big data from cheap phones', 'brain organoids', 'airborne networks', 'magnetic-resonance force microscopy', 'light-trapping photovoltaics', 'Reusable Rockets', 'bayesian machine learning', 'cancer genomics', 'nuclear reprogramming', 'atomic magnetometers']
Cluster5	['genome editing', 'solar fuel', 'reality mining', 'apple pay', 't-rays', 'nanowires', 'intelligent software assistant', 'neuron control', 'brain mapping', 'grid computing', 'modeling surprise', 'nanimprint lithogrpahy', 'connectomics', 'brain machine interface', 'engineered stem cells', 'Conversational Interfaces', 'data mining', 'biological machines', 'invisible revolution', '$100 genome', 'light field photography', 'natural language processing', 'diffusion tensor imaging', 'power grid control', 'glycomics']

Cluster	Technologies
Cluster6	['supergrids', 'temporary social media', 'cognitive radio', 'augmented reality', 'wireless power', 'crash-proof code', 'microfluidic optical fibers', 'separating chromosomes', 'pervasive wireless', 'facebook timeline', 'robot design', 'comparative interactomics', 'deep learning', 'digital imaging reimagined', 'high speed material discovery', 'flexible transistors', 'a new focus for light', 'nanomedicine', 'internet dna', 'mobile collaboration', 'oculus rift', 'social tv', 'agile robots', 'stretchable silicon', 'mechatronics']
Cluster7	['gestural interfaces', 'liquid biopsy', 'cellulolytic Enzymes', 'rNAi interference', 'universal authentication', 'wireless sensor networks']

Continuing with Example 3, described in the previous section and considering the pair of technologies ('magic leap','solar fuel), we can see that the weighted shortest path length is 3.5 with 4 links instead of a length of 5 and 5 links as in the technology concept network. The output of the program is shown below with Table 4.23 summarizing the weights of the links. The output of the program is shown below and verifies the existence of a path between the two nodes, provides the length of its shortest path, the number of links as well as the non-unique paths and the recommended path shown on Figure 4-10.

- There is a path from source to target? True
- shortest path lengths in the graph: 3.5
- Automatically generated shortest path in the graph: ['magic leap', 'nanomedicine', 'a new focus for light', 'brain machine interface', 'solar fuel']
- Number of links in the shortest path: 4
- Possible path: ['magic leap', 'nanomedicine', 'a new focus for light', 'brain machine interface', 'solar fuel']
- Number of distinct shortest paths in the graph: 1
- Recommended Path(s):['magic leap', 'nanomedicine', 'a new focus for light', 'brain machine interface', 'solar fuel']

Table 4.23: Links' Weights between 'magic leap' and 'solar fuel'- Keywords Network

Source	Target	Weighted Length
magic leap	nanomedicine	1.0
brain machine interface	solar fuel	1.0
a new focus for light	nanomedicine	1.0
brain machine interface	a new focus for light	0.5

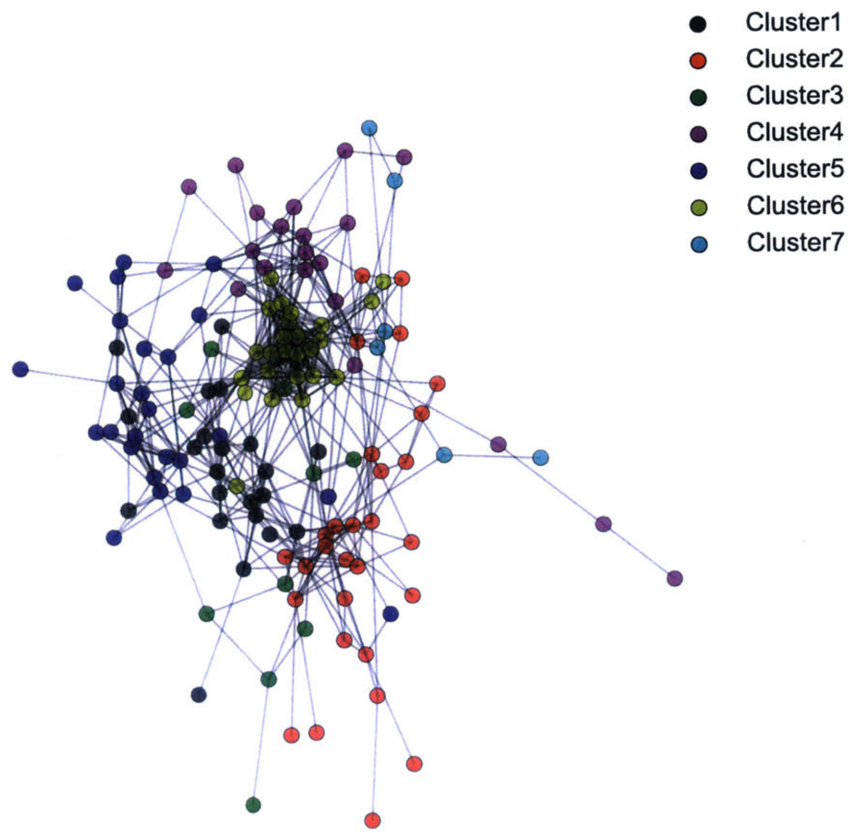

Figure 4-9: Clusterization for the Keywords-based Technology Network

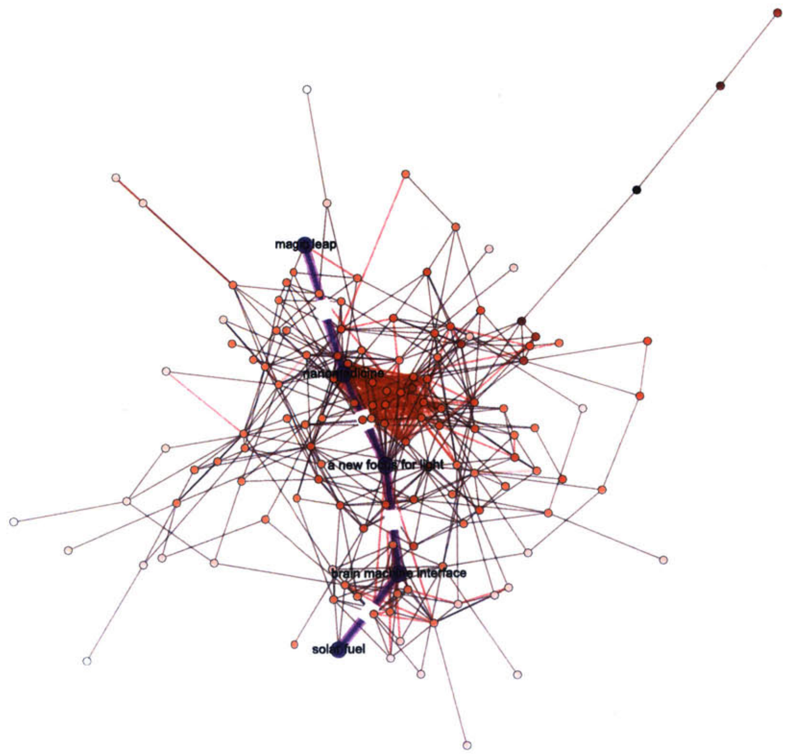

Figure 4-10: Shortest Paths between 'magic leap' and 'solar fuel' - Keywords

3. Technology linkage through Taxonomy Taxonomy represents the pre-defined technology domains that IBM AlchemyAPI tool uses to classify the technologies. The metrics of the network show a very high number of the links 8417 (an order of magnitude higher than the technology concepts network). The average degree is also an order of magnitude higher than that of the technology concepts network and there is only one connected component that includes all the nodes in the network. Table 4.24 summarizes the network's metrics.

Table 4.24: Network Metrics for the Taxonomy-based Technology Network

Measure	Value
Number of nodes	150

Measure	Value
Number of edges	8417
Average degree	112.22
Number of connected components	1
Size of largest connected component	150
The average shortest path length	1.24

These results seem to indicate that all the technologies are linked to each other, which is less informative than the two previous dimensions analyzed (concepts and keywords).

Table 4.25 summarizes for the centrality measures derived for the first five technologies.

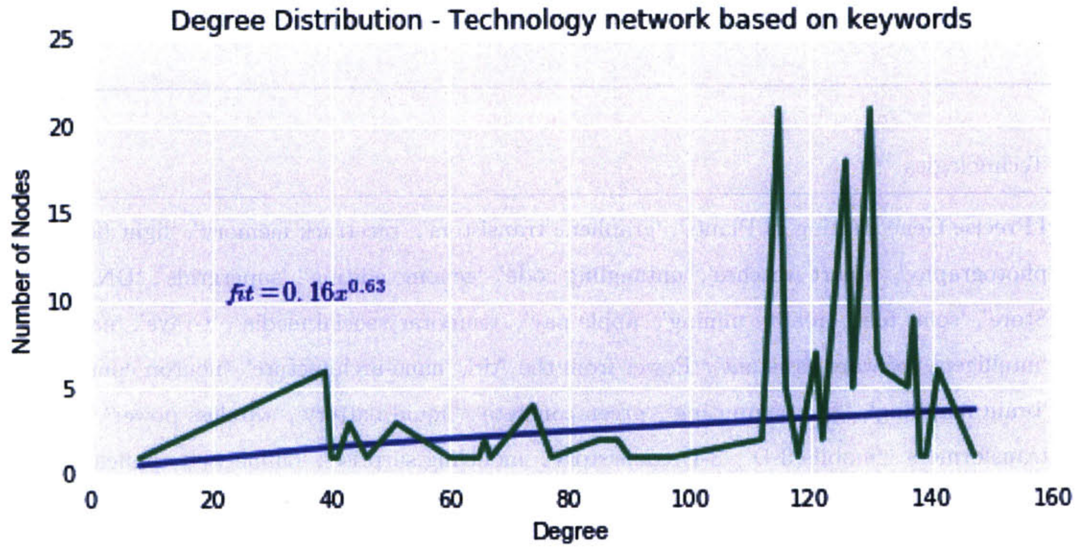

Figure 4-11: Nodes' Degree Distribution for the taxonomy-based Technology Network

Table 4.25: Nodes' Centrality Measures for the Taxonomy-based Technology Network

node	degree	betweenness	closeness	score
Precise Gene Editing in Plants	115	0.000000	0.814208	115.814208
graphene transistors	121	0.000686	0.841808	121.842494
racetrack memory	141	0.005902	0.949045	141.954947
supercharged photosynthesis	147	0.007709	0.986755	147.994464
light field photography	132	0.002028	0.897590	132.899618

The power law fit of the degree distribution shows results in a negative power parameter of -0.63 compared to the 0.43 from the technology concepts network. Figure 4-11 shows the degree distribution and the power fit curve.

The clusterization algorithm identifies three communities summarized on Table 4.26 and shown on Figure 4-12. Similarly to the keywords case we can see a mix of technology domains within a given cluster.

Table 4.26: Clusterization of the Taxonomy-based Technology Network

Cluster	Technologies
Cluster 1	['Precise Gene Editing in Plants', 'graphene transistors', 'racetrack memory', 'light field photography', 'smart watches', 'untangling code', 'genome editing', 'supergrids', 'DNA App Store', 'solar fuel', 'reality mining', 'apple pay', 'temporary social media', 't-rays', 'nanowires', 'intelligent software assistant', 'Power from the Air', 'nano-architecture', 'neuron control', 'brain mapping', 'grid computing', 'green concrete', 'liquid battery', 'wireless power', 'smart transformers', 'mobile 3-D', '3-D transistors', 'modeling surprise', 'offline web applications', 'microscale 3-D printing', 'liquid biopsy', 'separating chromosomes', 'probabilisitic chips', 'metabolomics', 'cellulolytic Enzymes', 'biomechanics', 'traveling wave reactor', 'SolarCity Gigafactory', 'microfluidics', 'baxter the blue collar robot', 'nanimprint lithogrpahy', 'epigenetics', 'robot design', 'deep learning', 'brain machine interface', 'nanosolar cells', 'bacterial factories', 'data mining', 'rNAi interference', 'software defined networking', 'digital imaging reimagined', 'quantum wires', 'prenatal DNA sequencing', 'high speed material discovery', 'cell-phone viruses', 'big data from cheap phones', 'brain organoids', 'a new focus for light', '$100 genome', 'internet dna', 'real-time search', 'airborne networks', 'injectible tissue engineering', 'universal authentication', 'synthetic cells', 'ultraprivate smartphones', 'Reusable Rockets', 'implantable electronics', 'oculus rift', 'bayesian machine learning', 'natural language processing', 'social tv', 'diffusion tensor imaging', 'agile robots', 'atomic magnetometers', 'stretchable silicon', 'mechatronics', 'project Loon', 'personal genomics']
Cluster 2	['supercharged photosynthesis', 'dual-action antibodies', 'personalized medical monitors', 'social indexing', 'hashCache', 'cognitive radio', 'enviromatics', 'ultra-efficient solar power 1', 'crash-proof code', 'crowdfunding', 'car to car communication', 'microfluidic optical fibers', 'nanopore sequencing', 'egg stem cells', 'pervasive wireless', 'facebook timeline', 'distributed storage', 'molecular imaging', 'silicon photonics', 'Robots That Teach Each Other', 'solid state batteries', 'Conversational Interfaces', 'nanobiomechanics', 'solar micro-grids', 'smart wind and solar', 'Immune Engineering', 'nanohealing', 'nanocharging solar', 'digital rights management', 'invisible revolution', 'nanomedicine', 'magnetic-resonance force microscopy', 'light-trapping photovoltaics', 'agricultural drones', 'wireless sensor networks', 'universal memory', 'software assurance', 'cancer genomics', 'power grid control', 'additive manufacturing', 'quantum cryptogrpahy']

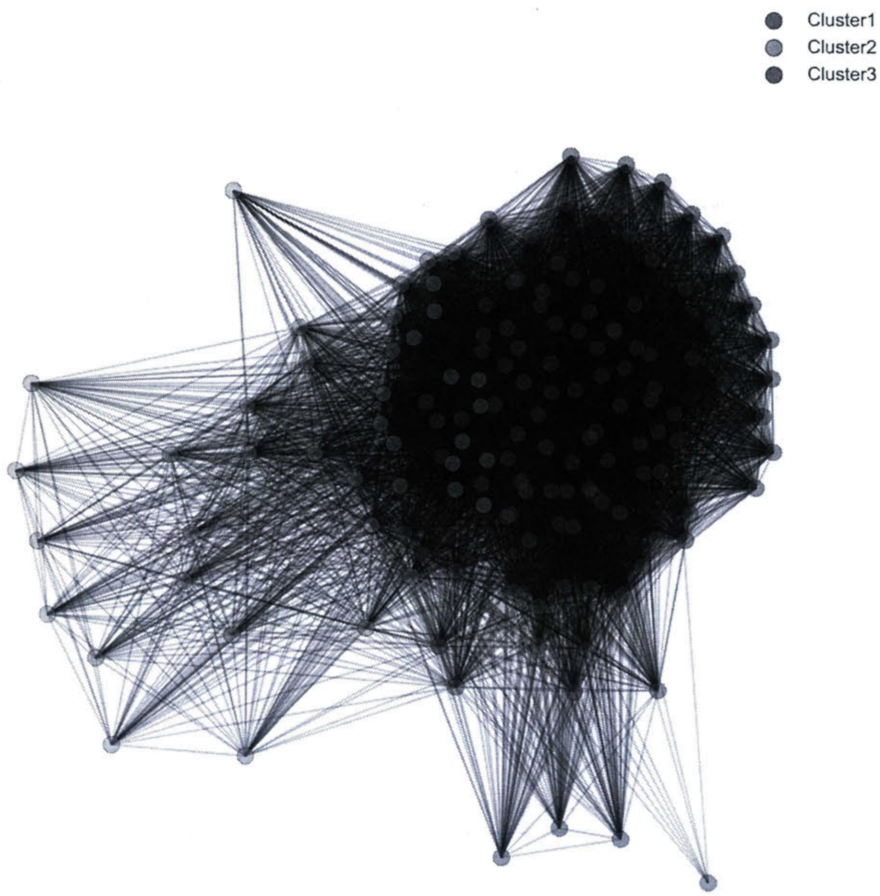

Figure 4-12: Clusterization for the Taxonomy-based Technology Network

Cluster	Technologies
Cluster3	['cloud streaming', 'Slack', 'paper diagnostics', 'gestural interfaces', 'cloud programming', 'augmented reality', 'single-cell analysis', 'megascale desalination', 'homomorphic encryption', 'comparative interactomics', 'nanopiezoelectronics', 'nanoradio', 'engineered stem cells', 'connectomics', 'neuromorphic chips', 'ultra-efficient solar power 2', 'microphotonics', 'biological machines', 'memory implants', 'flexible transistors', "peering into video's future", 'biometrics', 'universal translation', 'mobile collaboration', 'Tesla Autopilot', 'a faster Fourier Transform', 'synthetic biology', 'glycomics', 'nuclear reprogramming', 'magic leap']

To understand how this linkage dimension impacts the paths between technology, we analyze the same example (Example 3 from the technology concepts network) linking 'magic leap' and 'solar fuel'. The output of the program is summarized below including the weights of the individual links along the shortest path shown on Table 4.27. Figure 4-13 shows the recommended path on the network.

- There is a path from source to target? True
- shortest path lengths in the graph: 0.583333333333
- automatically generated shortest path in the graph: ['magic leap', 'Tesla Autopilot', 'solar fuel']
- Number of links in the shortest path: 2
- Possible path: ['magic leap', 'Tesla Autopilot', 'solar fuel']
- Number of distinct shortest paths in the graph: 1
- Recommended Path(s):['magic leap', 'Tesla Autopilot', 'solar fuel']

Table 4.27: Links' Weights between 'magic leap' and 'solar fuel'- Taxonomy Network

Source	Target	Weighted Length
magic leap	Tesla Autopilot	0.333333
Tesla Autopilot	solar fuel	0.250000

4. Technology linkage through Entities* (companies)

In this network dimension we explore how technologies are connected based on the companies they have links to. As can be seen from Table 4.28, the network is highly connected with more links than the technology concepts network a higher average degree and all but one technology within a single connected component.

Table 4.28: Network Metrics for the Entities-based Technology Network

Measure	Value
Number of nodes	150
Number of edges	2584
Average degree	34.45
Number of connected components	2
Size of largest connected component	149
The average shortest path length	1.81

Table 4.29 summarizes the centrality measure for this network and the Figure 4-14 shows the power law

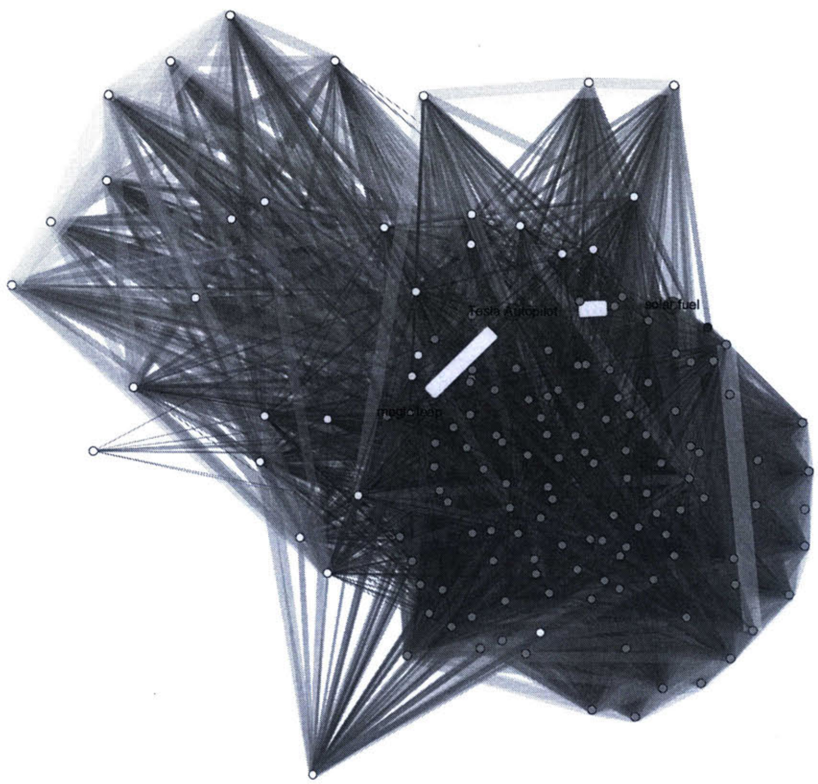

Figure 4-13: Shortest Paths between 'magic leap' and 'solar fuel' - Taxonomy

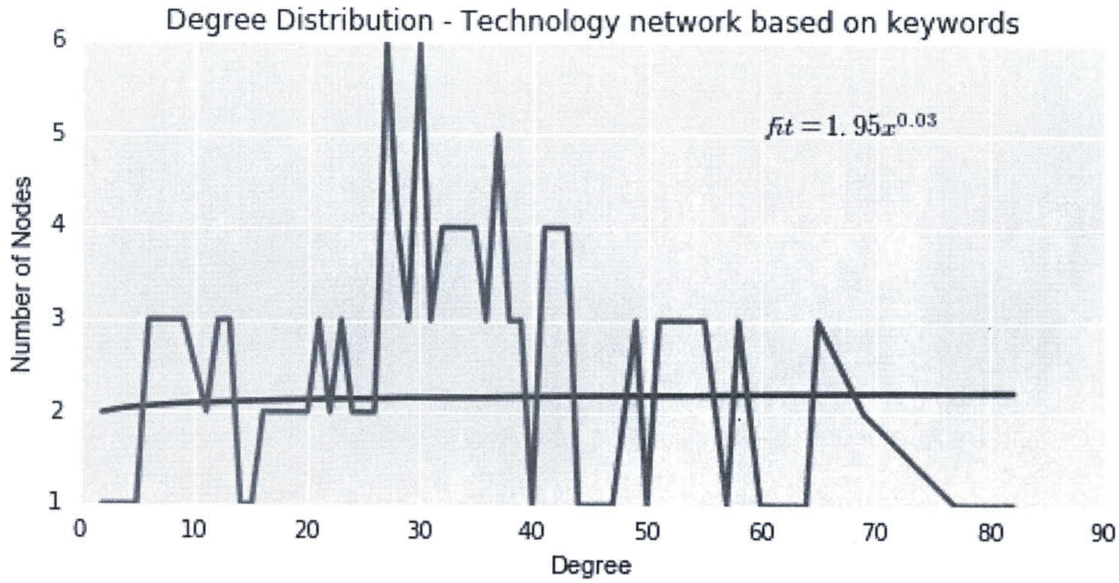

Figure 4-14: Nodes' Degree Distribution for the Entities-based Technology Network

fit of the degree distribution. Similarly to the technology taxonomy network, the power law parameter is negative (-0.03), an indication of a poor fit to this data set.

Table 4.29: Nodes' Centrality Measures for the Entities-based Technology Network

node	degree	betweenness	closeness	score
ultraprivate smartphones	32	0.001407	0.554307	32.555714
Precise Gene Editing in Plants	6	0.000029	0.458204	6.458234
graphene transistors	12	0.001411	0.501695	12.503106
racetrack memory	24	0.000886	0.530466	24.531352
grid computing	49	0.012870	0.599190	49.612060

The clustering algorithm identifies six communities described in Table 4.30 and shown on Figure 4-15. Note that this clusterization is not a measure of similarity unlike the previous dimensions analyzed (concepts, keywords and taxonomy) this dimension is indication of how companies build portfolios that span different technology domains.

Table 4.30: Clusterization of the Entities-based Technology Network

Cluster	Technologies
Cluster1	[‘Precise Gene Editing in Plants’, ‘graphene transistors’, ‘biometrics’, ‘personalized medical monitors’, ‘solar fuel’, ‘hashCache’, ‘paper diagnostics’, ‘cognitive radio’, ‘nano-architecture’, ‘cloud programming’, ‘data mining’, ‘modeling surprise’, ‘smart watches’, ‘microscale 3-D printing’, ‘microfluidics’, ‘distributed storage’, ‘homomorphic encryption’, ‘nanoradio’, ‘digital imaging reimagined’, ‘high speed material discovery’, ‘software defined networking’, ‘magnetic-resonance force microscopy’, ‘injectible tissue engineering’, ‘mobile collaboration’, ‘Tesla Autopilot’, ‘ultraprivate smartphones’, ‘implantable electronics’, ‘atomic magnetometers’, ‘mechatronics’, ‘project Loon’]
Cluster2	[‘racetrack memory’, ‘reality mining’, ‘apple pay’, ‘Slack’, ‘nanowires’, ‘intelligent software assistant’, ‘green concrete’, ‘augmented reality’, ‘wireless power’, ‘smart transformers’, ‘3-D transistors’, ‘crowdfunding’, ‘car to car communication’, ‘microfluidic optical fibers’, ‘separating chromosomes’, ‘probabilisitic chips’, ‘biomechanics’, ‘traveling wave reactor’, ‘SolarCity Gigafactory’, ‘nanimprint lithogrpahy’, ‘epigenetics’, ‘Robots That Teach Each Other’, ‘Conversational Interfaces’, ‘nanohealing’, ‘neuromorphic chips’, ‘flexible transistors’, ‘microphotonics’, ‘digital rights management’, ‘invisible revolution’, ‘a new focus for light’, ‘internet dna’, ‘light-trapping photovoltaics’, ‘synthetic cells’, ‘light field photography’, ‘natural language processing’, ‘universal memory’, ‘cancer genomics’, ‘diffusion tensor imaging’, ‘crash-proof code’, ‘additive manufacturing’]
Cluster3	[‘grid computing’, ‘bacterial factories’, ‘enviromatics’, ‘supercharged photosynthesis’, ‘mobile 3-D’, ‘liquid biopsy’, ‘metabolomics’, ‘nanopore sequencing’, ‘pervasive wireless’, ‘baxter the blue collar robot’, ‘silicon photonics’, ‘nanopiezoelectronics’, ‘brain machine interface’, ‘airborne networks’, ‘Reusable Rockets’, ‘a faster Fourier Transform’, ‘bayesian machine learning’, ‘glycomics’, ‘magic leap’]
Cluster4	[‘untangling code’, ‘supergrids’, ‘social indexing’, ‘temporary social media’, ‘t-rays’, ‘Power from the Air’, ‘neuron control’, ‘brain mapping’, ‘facebook timeline’, ‘molecular imaging’, ‘connectomics’, ‘solid state batteries’, ‘nanosolar cells’, ‘megascale desalination’, ‘nanobiomechanics’, ‘smart wind and solar’, ‘ultra-efficient solar power 2’, ‘biological machines’, ‘nanomedicine’, ‘real-time search’, ‘universal translation’, ‘oculus rift’, ‘social tv’, ‘software assurance’, ‘power grid control’, ‘stretchable silicon’]

Cluster	Technologies
Cluster5	['dual-action antibodies', 'cloud streaming', 'gestural interfaces', 'single-cell analysis', 'liquid battery', 'egg stem cells', 'Immune Engineering', 'big data from cheap phones', '$100 genome', "peering into video's future", 'universal authentication', 'agile robots', 'nuclear reprogramming']
Cluster6	['genome editing', 'DNA App Store', 'ultra-efficient solar power 1', 'offline web applications', 'cellulolytic Enzymes', 'robot design', 'comparative interactomics', 'engineered stem cells', 'agricultural drones', 'solar micro-grids', 'rNAi interference', 'deep learning', 'nanocharging solar', 'prenatal DNA sequencing', 'memory implants', 'cell-phone viruses', 'brain organoids', 'synthetic biology', 'wireless sensor networks', 'personal genomics', 'quantum cryptogrpahy']

Continuing with the same technology pair ('magic leap','solar fuel') as the source and the target respectively, the shortest path length is estimated to be 1.33 with three links.

The shortest path returned by the program is ['magic leap', 'ultra-efficient solar power 1', 'distributed storage', 'solar fuel'], the weights of each of the links in this path are summarized in Table 4.31.

Table 4.31: Links' Weights between 'magic leap' and 'solar fuel'- Entities Network

Source	Target	Weighted Length
magic leap	ultra-efficient solar power 1	0.500000
distributed storage	ultra-efficient solar power 1	0.500000
distributed storage	solar fuel	0.333333

Based on the comparison of the four dimensions for technology links explored in this section (concepts, keywords, taxonomy and companies) we conclude that concepts are the most suitable dimension. The technology concepts network will be used in the remainder of the analysis.

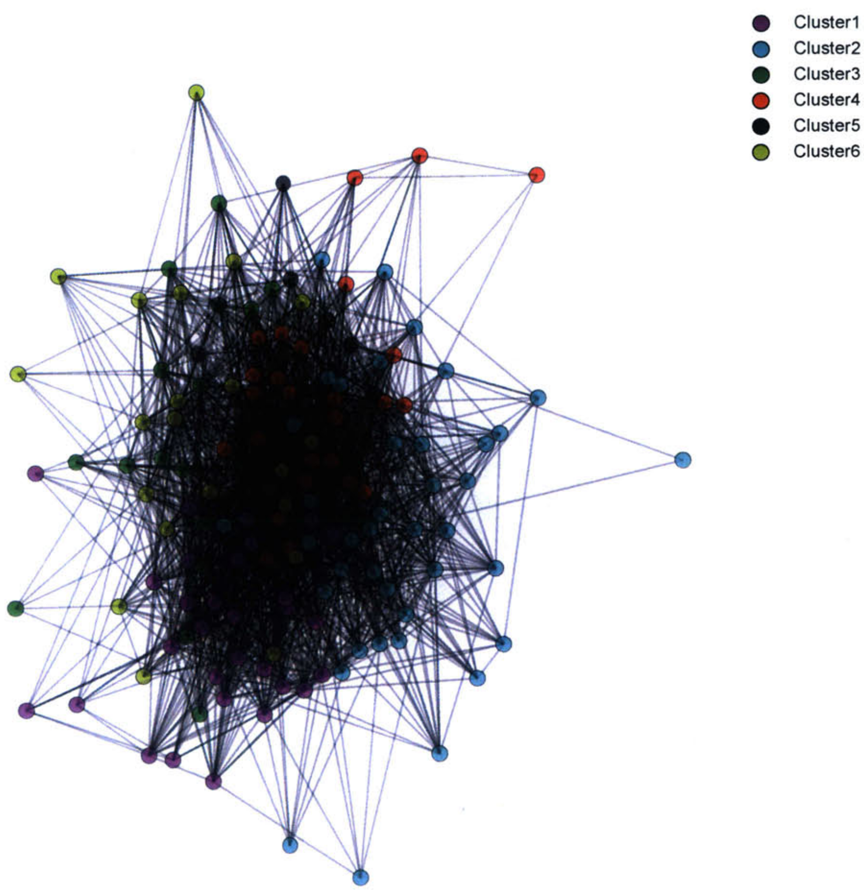

Figure 4-15: Clusterization for the Entities-based Technology Network

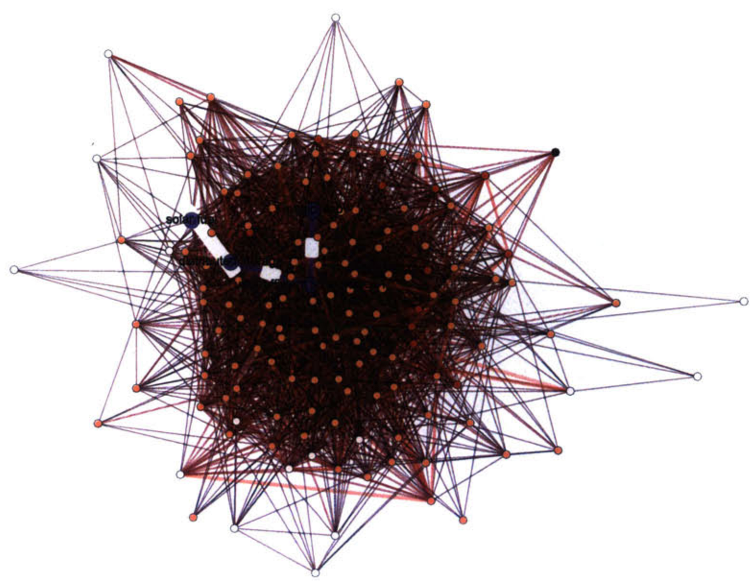

Figure 4-16: Shortest Paths between 'magic leap' and 'solar fuel' - Entities

4.2.2 Company to Company Networks

In this section we explore the linkages between companies using two dimensions that are particularly relevant to uncovering the linkages between companies. The first dimension is "concepts", as seen in the previous section, concepts represent the most suitable relationship linkage between technologies. The second dimension is to link companies based on the technologies they share. This second dimension reveals a technology based companies relationships that will be particularly interesting when analyzing company strategies.

Company linkage through Concepts

Table 4.32 summarizes the metrics of the network of companies based on concepts. When compared to the concepts technology network discussed in Section 1, we note that there are twice the number of nodes and close to four times the number of links. The average degree is high 35.5 and the majority of nodes are connected to a single large component. The nodes' centrality measures are presented in Table 4.32.

Table 4.32: Network Metrics for the Concepts-based Company Network

Measure	Value
Number of nodes	229
Number of edges	4067
Average degree	35.51
Number of connected components	4
Size of largest connected component	224
The average shortest path length	2.02

Table 4.33: Nodes' Centrality Measures for the Concepts-based Company Network

user	degree	betweenness	closeness	score
Amgen	22	0.000714	0.507973	22.508687
Roche	54	0.001502	0.542579	54.544081
First Solar	16	0.000194	0.497768	16.497962
Airbnb	45	0.000087	0.527187	45.527274
Alta Devices	8	0.001008	0.371048	8.372056

The power law fit of the degree distribution is shown on Figure 4-17. It has a parameter of 0.22 and the distribution shows a long tail with two degree ranges dominating it; the first degree range is between 5 and 20 and the second degree range is between 40 and 65.

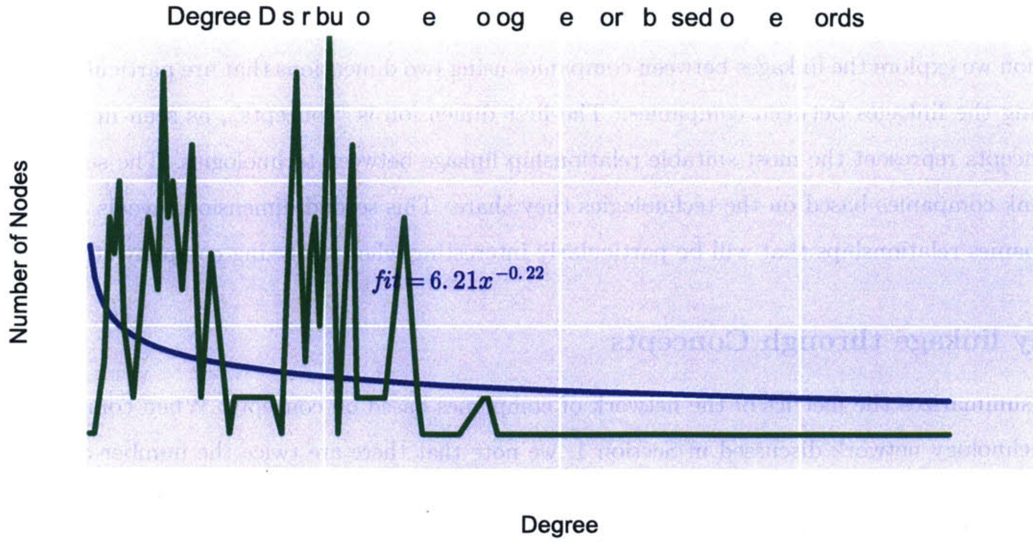

Figure 4-17: Nodes' Degree Distribution for the Concept-based Company Network

There are six communities extracted from the graph and the individual clusters include companies from different sectors as shown in Figure 4-18. The companies' communities are particularly interesting when considering competitive games as they help define the sub-set of companies that could potentially compete against each other for a given technology. They are summarized in Table 4.34.

Table 4.34: Clusterization of the Concept-based Company Network

Cluster	Technologies
Cluster 1	['DuPont Pioneer', 'Napster', 'Abaron', 'Microgrids', 'Cellectis', 'Adobe', 'NANOELECTRONICS Nanotubes', 'Boston Dynamics', 'Rice group', 'Novacem', 'Investment Technology Group', 'Amazon', 'Bell Labs', 'Silicon', 'AeroVironment', 'Nortel', 'FedEx', 'Quake', 'Ericsson', 'Wilson Sporting Goods', 'Helix', 'IDC', 'Baidu', 'Cytel', 'CacheLogic', 'Warburg Pincus', 'Oculus', 'Epigenomics', 'NREL', 'InnoPath Software', 'Kinect', 'Autodesk', 'Tesla', 'Johns Hopkins', 'Lucent Technologies', 'F-Secure', 'Pioneer', 'Phenomenome Discoveries', 'TELECOM Wireless', 'Mera Gao Power', 'ValleZ', 'Brain Corp.', 'Netflix', 'Joule', 'CFM International', 'National Public Radio', 'Double Fine Productions']

Table 4.34: Clusterization of the Concept-based Company Network

Cluster	Technologies
Cluster2	['Greer designs', 'Modern Physics', 'DaimlerChrysler', 'Honda', 'Zrich', 'E Ink', 'DIY Drones', 'Flagship Ventures', 'Digital Biometrics', 'OvaScience', 'Nantero', 'Genentech', 'BMW', 'Hitachi', 'Paulson', 'Suntech Power', 'TBD', 'Wildcat Discovery Technologies', 'Grant Willson', 'Semprius', 'NTT DoCoMo', 'Cadillac', 'IDE Technologies', 'Twitter', 'MRFM', 'OpenFlow', 'OnLive', 'Pfizer', 'RoboBrain', 'BioNanomatrix', 'Google', '3D Robotics', 'Japanese', 'Cisco', 'Vertex Pharmaceuticals', 'McAfee']
Cluster3	['Rice University', 'Airbnb', 'Progenics Pharmaceuticals', 'AspectJ', 'Isermann', 'YouTube', 'Starbucks', 'FriendFeed', 'Cryptography Research', 'Intellectual Ventures', 'Warner Brothers', 'GE', 'Whole Foods Market', 'BT', 'Checkfree', 'eBay', 'Samsung', 'Facebook', 'Patricia Price', 'Sony', 'Globus Toolkit', 'Nokia', 'Magic Leap', 'Artivest', 'Verizon', 'SolarCity', 'AM radio', 'IRRI', 'NEC', 'Carnegie Mellon', 'Associated Press', 'Juno', 'VPL Research', 'Toshiba', 'SpaceX', 'TinyOS', 'Snapchat', 'PricewaterhouseCoopers', 'Synthetic Genomics', 'DuEr', 'PayPal', 'Siri', 'DisplaySearch', 'TeraGrid', 'GlaxoSmithKline', 'Cochlear', 'Great Ormond', 'Wal-Mart', 'Pandora', 'CRISPR', 'Eagle', 'SRI', 'Visionics', 'Microsoft', 'Perlegen Sciences', 'Sprint']
Cluster4	['Israel Desalination Enterprises', 'Amgen', 'exabytes', 'Palm Pilots', 'Xerox', 'Hulu', 'Agilent Technologies', 'U. Oklahoma', 'Oxford', 'Asus', 'Kitching', 'Caribou Biosciences', 'Accenture', 'TALENs', 'Counterpane', 'Xcel Energy', 'HoloLens', 'Cellular Dynamics', 'Palm Computing', 'Motorola', 'MindNet', 'Miniaturizing radios', 'Silent Circle', 'SunPower', 'Advanced Cell Technology', 'Roche', 'Pebble', 'Philips Research', 'RAM', 'Mycometrix', 'Alzheimer', 'Meta Group', 'MatchMaker Exchange', 'Kazaa', 'Bing', 'Adobe Systems', 'CovX', 'HashCache', 'Juniper', 'Panasonic', 'IBM', 'Veritas Genetics', 'Abaron Biosciences', 'Intel', 'Apple', 'RunKeeper', 'ABB', 'Forrester Research', 'HBMIs', 'OneWorld Health', 'Hunch', 'ContentGuard', 'Seagate']
Cluster5	['U. Michigan', 'First Solar', 'Wired', 'Merck', 'GM', 'Orange', 'Digital Equipment', 'mycoplasmas', 'Cell Design Labs', 'Rethink Robotics', 'Mako Surgical', 'Silevo', 'Fujitsu', 'Hewlett-Packard', 'Georgia Tech', 'GetGlue', 'Illumina', 'Poynt', 'Nintendo', 'Verinata', 'Continental', 'helium', 'Citibank', 'Toyota']
Cluster6	['Alta Devices', 'Blackphone', 'T-Mobile', 'Amyris Biotechnologies', 'nanoworld', 'Institute of Molecular Biotechnology', 'Siemens', 'Sirtris Pharmaceuticals']

To continue with the same example as in the technology network analysis section, we focus on the two

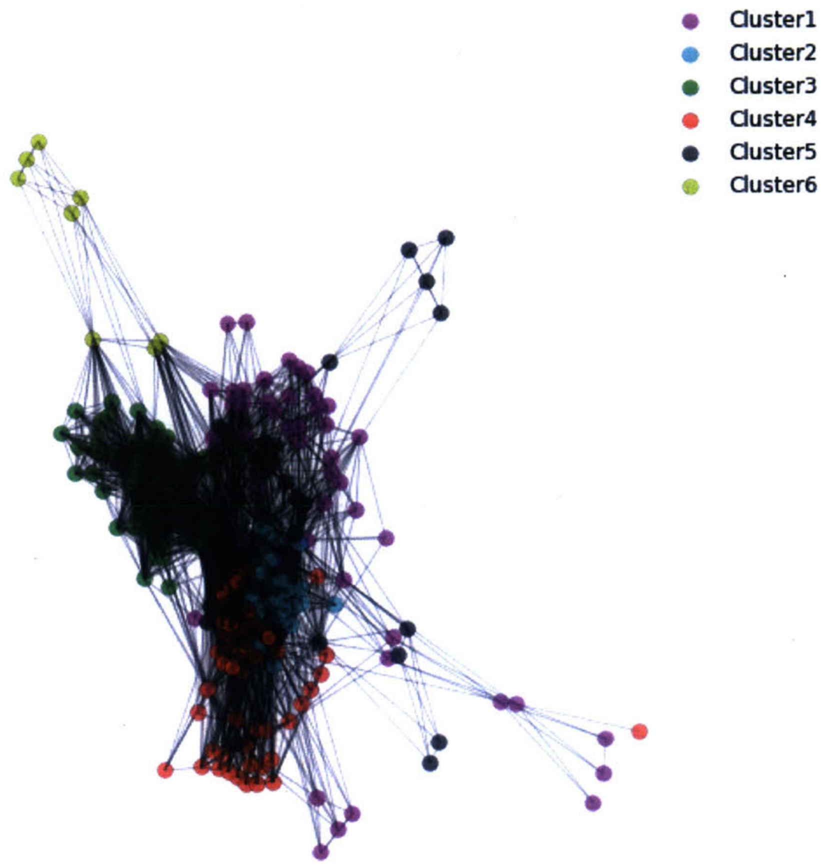

Figure 4-18: Clusterization for the Concept-based Technology Network

technologies 'magic leap' and 'solar fuel' by analyzing the two companies that own the technologies. These are 'Magic Leap' and 'Joule' respectively. The shortest path between the two is shown in Figure 4-19 and the paths' weights are summarized in Table 4.35.

- There is a path from source to target? True
- Shortest path lengths in the graph: 1.11298701299
- Automatically generated shortest path in the graph: ['Magic Leap', 'Rethink Robotics', 'CFM International', 'Joule']
- Number of links in the shortest path: 3
- Number of distinct shortest paths in the graph: 1

Table 4.35: Links' Weights between 'magic leap' and 'solar fuel'- Company and Concepts Network

Source	Target	Weighted Length
Joule	CFM International	1.000000
CFM International	Magic Leap	0.333333
Rethink Robotics	CFM International	0.012987
Rethink Robotics	Magic Leap	0.100000

2. Company linkage through Technology The second dimension of interest for linking companies is the along the technologies they are connected to. This shows technological similarities. On average each company is connected to 4.6 other companies, there are more clusters however than in any other network, 49. The metrics of the network are summarized in Table 4.36.

Table 4.36: Network Metrics for the Technology-based Company Network

Measure	Value
Number of nodes	229
Number of edges	525
Average degree	4.58
Number of connected components	49
Size of largest connected component	138
The average shortest path length	2.91

Figure 4-20 shows the histogram distribution of the connected components. In this section we will

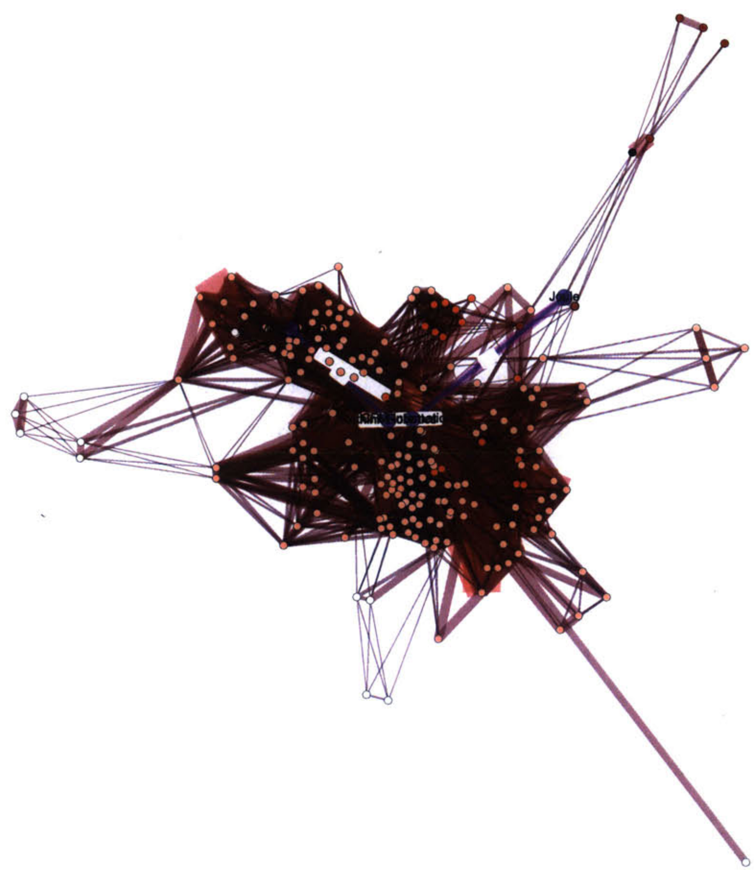

Figure 4-19: Shortest Paths between 'magic leap' and 'solar fuel' - Company and Concepts

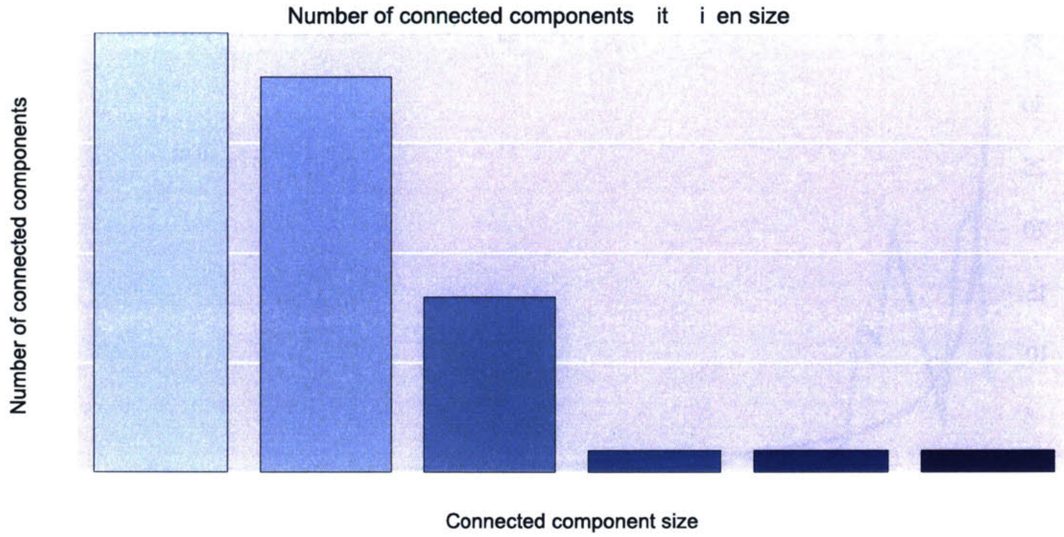

Figure 4-20: Histogram of Size and Number of Connected Components in the Technology-based Company Network

focus on the largest connected component of size 138 as the sizes of the other connected components starts decreasing rapidly.

Table 4.37 summarizes the centrality measures for the companies in the network.

Table 4.37: Nodes' Centrality Measures for the Technology-based Company Network

node	degree	betweenness	closeness	score
Greer designs	8	0	0.402941	8.402941
Israel Desalination Enterprises	4	0	0.309955	4.309955
MRFM	7	0	0.268627	7.268627
Continental	4	0	0.353093	4.353093
U. Michigan	3	0	0.274549	3.274549

Figure 4-21 shows the power law distribution of the degree distribution with a parameter of 0.93.

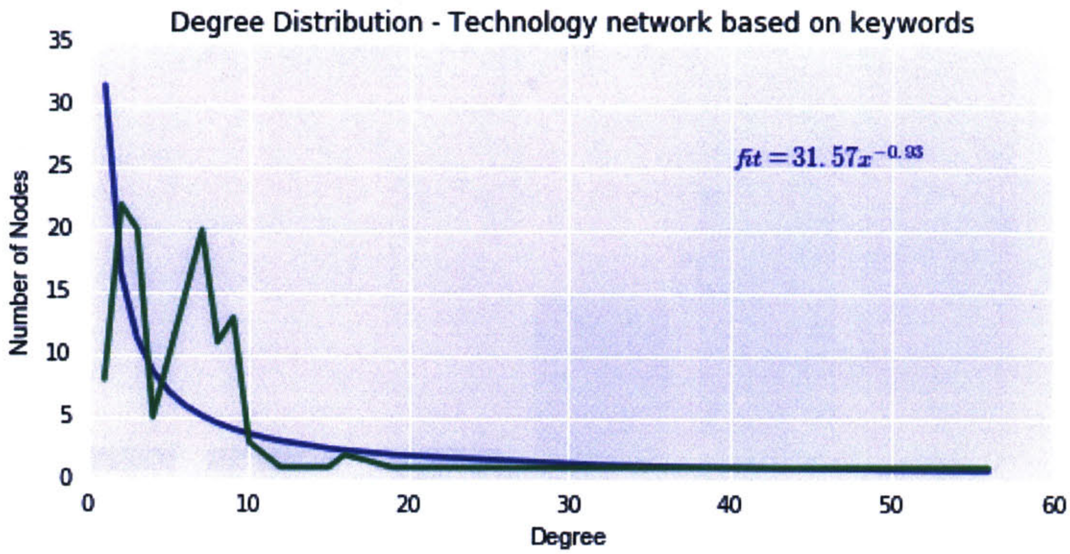

Figure 4-21: Nodes' Degree Distribution for the Technology-based Company Network

There are nine communities extracted from the network. When examining Table 4.38 we can see that these companies are from different technology domains. Figure 19 shows the none cluster in the network.

Table 4.38: Clusterization of he Technology-based Company Network

Cluster	Technologies
Cluster1	['Greer designs', 'Napster', 'VPL Research', 'Cellular Dynamics', 'SunPower', 'Autodesk', 'Philips Research', 'Apple', 'Synthetic Genomics', 'National Public Radio', 'Abaron Biosciences', 'Mera Gao Power', 'Isermann', 'RunKeeper', 'Toyota']
Cluster2	['Israel Desalination Enterprises', 'AspectJ', 'Intellectual Ventures', 'Caribou Biosciences', 'Whole Foods Market', 'CovX', 'Nantero', 'Xcel Energy', 'Orange', 'Facebook', 'Poynt', 'Silevo', 'Fujitsu', 'Bell Labs', 'FedEx', 'Quake', 'exabytes', 'Wilson Sporting Goods', 'Juno', 'Hewlett-Packard', 'IDE Technologies', 'Amyris Biotechnologies', 'Adobe Systems', 'Rice group', 'DuEr', 'SRI', 'Double Fine Productions']
Cluster3	['MRFM', 'Continental', 'RAM', 'F-Secure', 'Palm Computing', 'BMW', 'Mako Surgical', 'Magic Leap', 'Oculus', 'Agilent Technologies', 'Japanese', 'Cisco']
Cluster4	['U. Michigan', 'Joule', 'AM radio', 'Meta Group', 'TELECOM Wireless', 'OpenFlow', 'HBMIs', 'Sprint']
Cluster5	['OnLive', 'Microgrids', 'Patricia Price', 'Georgia Tech', 'Nintendo', 'Forrester Research', 'Panasonic', 'CFM International', 'PayPal', 'DisplaySearch', 'GlaxoSmithKline', 'Seagate']

Cluster	Technologies
Cluster6	['Verizon', 'Cryptography Research', 'Blackphone', 'Rethink Robotics', 'Artivest', 'Visionics', 'Vertex Pharmaceuticals', 'Lucent Technologies', 'Lytro']
Cluster7	['DaimlerChrysler', 'Siemens', 'DIY Drones', 'NANOELECTRONICS Nanotubes', 'Digital Equipment', 'Cell Design Labs', 'Miniaturizing radios', 'Baidu', 'Silicon', 'Qualcomm', 'IDC', 'Boston Dynamics', 'Snapchat', 'McAfee', 'Progenics Pharmaceuticals', 'Kitching', 'Google', 'Hunch']
Cluster8	['Xerox', 'Starbucks', 'Alzheimer', 'Adobe', 'Siri', 'TALENs', 'Counterpane', 'Genentech', 'T-Mobile', 'Sony', 'Suntech Power', 'IBM', 'NTT DoCoMo', 'Warburg Pincus', 'Verinata', 'Epigenomics', 'TinyOS', 'Pioneer', 'Oxford', 'Brain Corp.', 'Great Ormond', 'OneWorld Health', 'CRISPR', 'Institute of Molecular Biotechnology', '3D Robotics', 'TBD']
Cluster9	['Zrich', 'Pfizer', 'GM', 'Novacem', 'MindNet', 'Amazon', 'Ericsson', 'Helix', 'Cadillac', 'Johns Hopkins', 'Wildcat Discovery Technologies']

Continuing on with the same pair of companies ('Magic Leap' and 'Joule') the shortest path extracted from the network is ['Magic Leap', 'CRISPR', 'Meta Group', 'Joule'] and has a length of 3.0. with three links. The path is shown on Figure 4-23. The weights along the path are summarized in Table 4.39.

- There is a path from source to target? True
- Shortest path lengths in the graph: 3.0
- Automatically generated shortest path in the graph: ['Magic Leap', 'CRISPR', 'Meta Group', 'Joule']
- Number of links in the shortest path: 3

Table 4.39: Links' Weights between 'Magic Leap' and 'Joule'- Company and Technology Network

Source	Target	Weighted Length
Meta Group	Joule	1
Meta Group	CRISPR	1
Magic Leap	CRISPR	1

In this section we have analyzed company connections along two dimensions concepts and technologies. The two networks have some notable differences. The companies' networks using concepts and technologies have the following differences respectively: number of links 4000 versus 500 for the technology, an average degree of approximately 35 versus 4, a number of connected components of approximately 4 versus 49 and

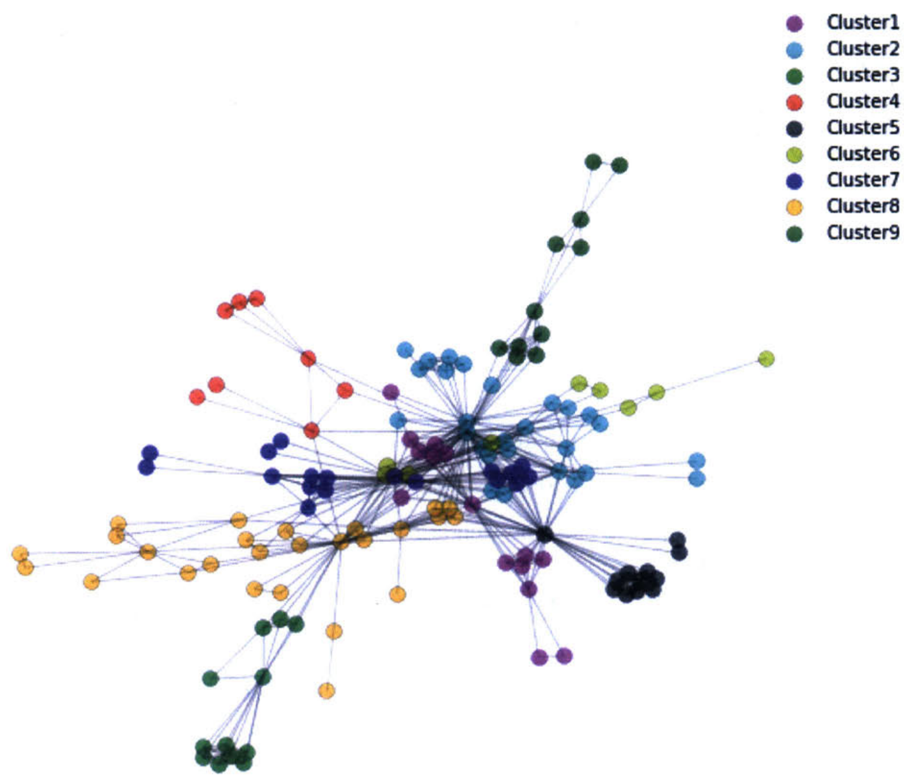

Figure 4-22: Clusterization for the Technology-based Company Network

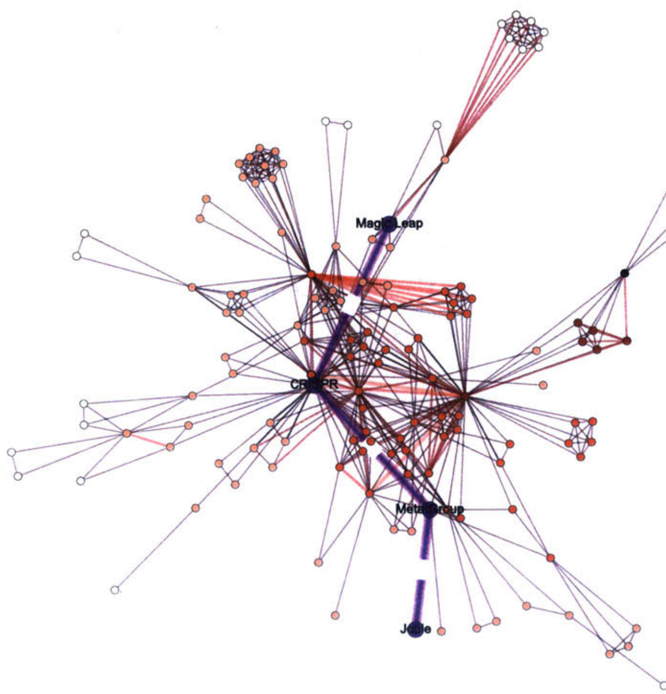

Figure 4-23: Shortest Paths between 'Magic Leap' and 'Joule'- Company and Technology Network

an average shortest path length of 2 versus almost 3.

Also when examining the communities extracted from the network, we suggest to use them in two ways. First, the communities extracted from the concepts can be used as an indication of similarity between the companies. Therefore, if a company from a given community is targeting a technology we can assume that companies from the same community would also be interested in the same technology. Second, for the communities extracted from the technology network, we can assume that two companies in the same community already have access to common technologies.

Combining the Companies and Technologies in a Single Network In this section, we examine the combined networks from technology and companies. The objective of this approach is to examine whether the recommended shortest path between two technologies and two companies are similar to those extracted from the networks independently. For this step we use the networks based on concepts.

Examining the path between the two companies 'Magic Leap' to 'Joule', and the path between the two technologies 'magic leap' and 'solar fuel' the combined network shows the following:

First, considering an unweighted graph, where all links have a value of 1, the path suggested is shown on Figure 4-24. We note that the path between the source technology and target technology in fact goes through the companies that have invested in these technologies.

- has a path from source to target?: Magic Leap to Joule True
- shortest path lengths in the graph: 2
- ['Magic Leap', 'CFM International', 'Joule']
- has a path from source to target?: magic leap to solar fuel True
- shortest path lengths in the graph: 3
- ['magic leap', 'Google', 'Flagship Ventures', 'solar fuel']

Second, considering a weighted graph. The weights of the intra-network links for the technologies and companies networks are inherited from the separate networks. The links between nodes from the two networks are set to the median value of 0.5. The shortest path length and recommended shortest path are highly sensitive to the weight attributed to the links between companies and technologies connected to them. Weights between 1 and 0.5 produce the shortest paths shown below on Figure 4-25, however lower weights introduce many intermediate nodes as it is less costly to move back and forth from the company layer to the technology layer.

- has a path from source to target?: Magic Leap to Joule True
- shortest path lengths in the graph: 1.11298701299
- ['Magic Leap', 'Rethink Robotics', 'CFM International', 'Joule']

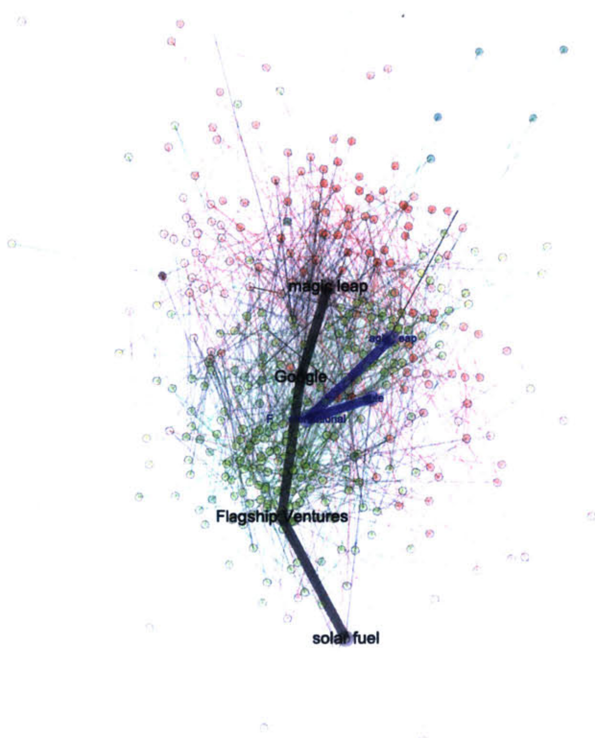

Figure 4-24: Shortest Paths between 'Magic Leap' and 'Joule'- Combined Company and Technology Network along Concepts - Unweighted Graph

- has a path from source to target?: magic leap to solar fuel True
- shortest path lengths in the graph: 2.15461255124
- ['magic leap', 'Microsoft', 'CFM International', 'Snapchat', 'Cochlear', 'Flagship Ventures', 'solar fuel']

Tradespace of possible paths

In this last section we generate a trade space from all the possible walks in the network between a pair of technologies (source, target). We focus on the technology network generated from concepts and define the benefit of each node and the cost of a path between two nodes. The benefit of a node is simply its score as defined in the previous section and the cost of a path is the weighted shortest length between a pair of nodes. Starting from all possible sources and all possible targets, we can generate the trade space of benefit versus cost. Table 4.40 summarizes the pairs of nodes with the highest cost, highest benefit and lowest cost.

Table 4.40: Cost and Benefit for all Pairs of Nodes in the Technology Network - Links represent Concepts

Characteristic	source	target	cost	benefit	name
highest cost	magic leap	megascale desalination	6.0	1.234323	magic leap_megascale desalination
highest cost	megascale desalination	magic leap	6.0	2.260738	megascale desalination_magic leap
highest benefit	augmented reality	smart watches	0.500000	32.52494	augmented reality_smart watches

Characteristic					
	source	target	cost	benefit	name
highest benefit	mobile 3-D	smart watches	0.500000	32.52494	mobile 3-D_smart watches
highest benefit	cell-phone viruses	smart watches	0.500000	32.52494	cell-phone viruses_smart watches
lowest cost	nanopore sequencing	100 dolalr genome	0.142857	27.362744	nanopore sequencing_$100 genome

Figure 4-26 shows the full trade-space as well as the nodes on the Pareto front.

For any given pair of source and target nodes in the trade space there could be non-unique paths in the network to link the two nodes.

To illustrate these non-unique paths, we took the pairs of nodes on the Pareto front and analyzed whether they exhibited multiple shortest paths options. This initial focus on the Pareto front is motivated by the fact these nodes represent the best options in the space for given costs.

Examining the trade space we can see that for a given benefit and cost there could be more than one pair of associated nodes. For instance, there are 9 node pairs on the Pareto front with a cost 0.14 of and a benefit of 23.35. These include: - ('separating chromosomes', 'internet dna') - ('separating chromosomes', 'personal genomics') - ('nanopore sequencing', 'personal genomics') - ('$100 genome', 'nanopore sequencing') - ('internet dna', 'separating chromosomes') - ('internet dna', 'personal genomics') - ('personal genomics', 'separating chromosomes') - ('personal genomics', 'nanopore sequencing') - ('personal genomics', 'internet dna')

All the pairs examined had a single shortest path each and mostly these paths are a single link between the two nodes. This result is not surprising given that the Pareto front will minimize cost which is represented by the length of the shortest path. If the path is represented by a single link and this link has the shortest

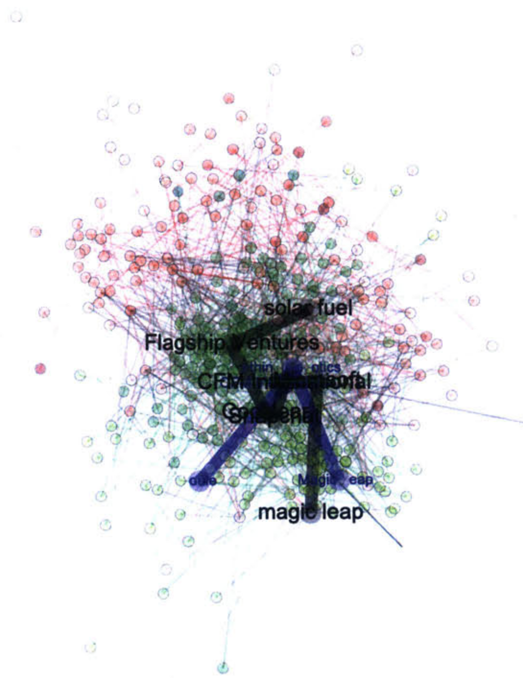

Figure 4-25: Shortest Paths between 'Magic Leap' and 'Joule'- Combined Company and Technology Network along Concepts - Weighted Graph

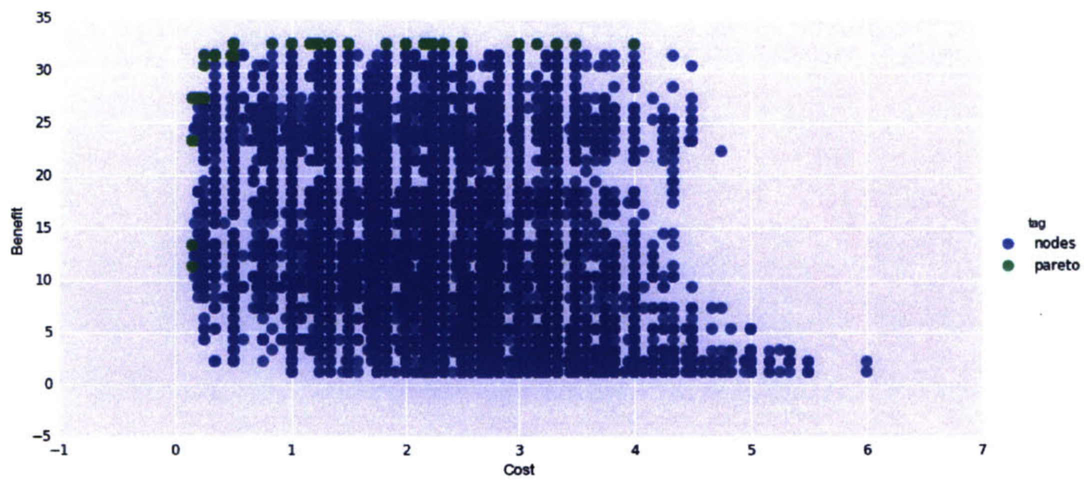

Figure 4-26: Benefit versus Cost Trade-space of all Nodes in the Technology Network - Linked by Concepts

weighted path calculated as the the inverse of the value from the adjacency matrix. It is less likely that there are multiple paths exhibiting the same characteristics. However as the number of segments in the path increase, and as the weighted length increase, it is more likely that other nodes and links would offer alternative paths.

Why is the trade space useful?

The trade space represents all the technology pairs and paths to link them. These represent all the options available to all the companies. One of the major advantages of the trade space is to be able to identify the most attractive options that a single firm has. It can also be used to examine whether node pairs in the same communities are more attractive than those across communities.

Figure 4-27, shows the position on the trade space of pairs of nodes belonging to different clusters generated in Section 1 (Technologies network based on concepts. The different colors represent the pairs (source, target) that belong to the same cluster. For instance the green represents all the possible combinations of nodes in Cluster 1 (Living Matter), whereas red represents the pairs in Cluster 2 (Energy). If nodes in a pair belong to different clusters the pair would appear in blue. We note that pairs of nodes within the (Living Matter) cluster have higher benefit and lower cost than those in energy for example.

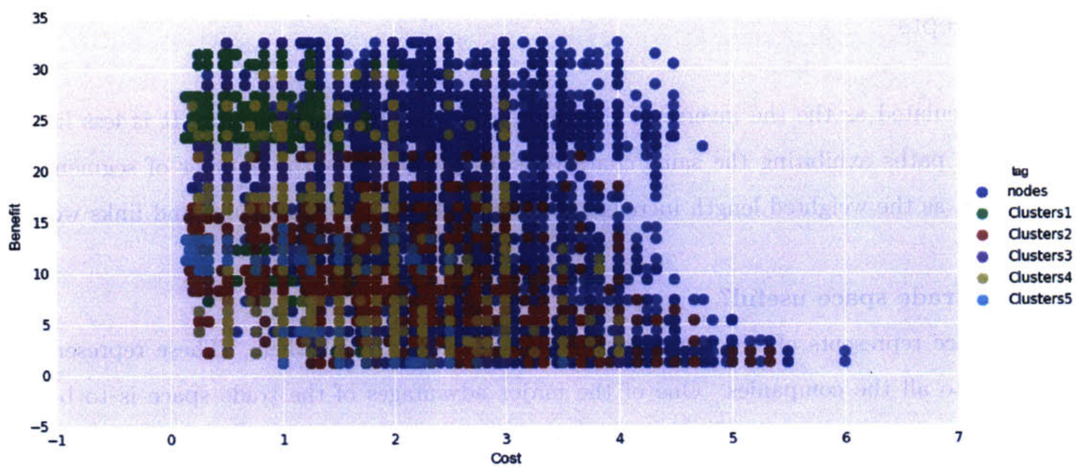

Figure 4-27: Benefit versus Cost Trade-space of all Nodes in the Technology Network - Linked by Concepts - Shown by Cluster

Chapter 5

Strategic Games

In this chapter we consider strategic decisions in which two or three significant participants interact with each other. This interaction becomes a strategic game when the participants become tied to each other and committed to the particular decision or transaction in which they are involved. As described in **Chapter 2**, the analysis of the outcomes of strategic games relies on the calculation of payoffs for the considered strategies. This is where the network and its metrics developed in **Chapter 4** come in to play as they allow us to select, for a given firm and a given target technology, the subset of feasible strategies among the trade-space of all possible strategies. Second they allow us to calculate the payoffs of the game. We limit the variables of the strategic game to the two competing companies and the target technology. The model developed in **Chapter 4** provides, for each one of the competitors, the best strategy including the starting technology and the network path leading to the target technology. It also provides, the payoff matrix of the game base don which we can identify dominating and dominated strategies if they exist.

A review of the networks developed in **Chapter 4** indicated that the concepts-based network offered the most relevant information regarding technology similarities. These networks allow us to simulate different paths in the technology landscape from a starting technology to a target technology. The starting technology for a given firm is one that belongs to its portfolio, whereas the target technology is one that the firm is considering to acquire. We note that in this case a technology belongs to a firm, if it has a connection to the firm but the firm does not necessarily has to own it. The structure of this chapter is as follows; first we will summarize the important network metrics from **Chapter 4**. Second we will identify the a firm's position in the network based on its technology portfolio. Third, we will define a number of games that can be played between at least two firms and show how the network metrics developed in **Chapter 4** help us compare different strategies. Finally, in the last section we will apply the framework to six use cases and analyze real-world competitive scenarios.

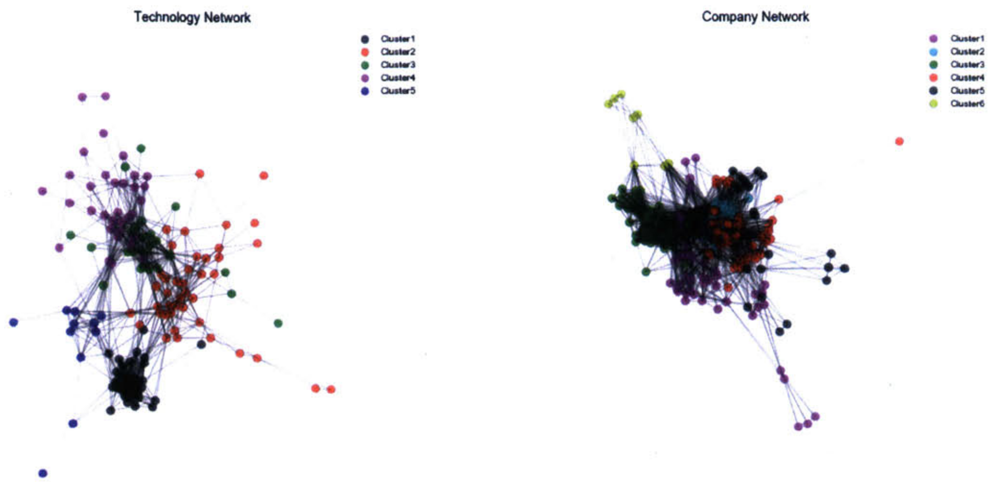

Figure 5-1: Technology Network (left) and Company Network (right) - with Internal Links Representing Concepts

5.1 Important network metrics for strategic games

We will start by summarizing the most important results from **Chapter 4**, starting with both networks representing technologies and companies each linked by concepts. Table 5.1 shows the metrics of the two networks while Figure 5-1 shows the networks' clusterization. The colors indicate that the nodes belong to the same cluster within a given network.

Table 5.1: Summary of Network Metrics for Technologies and Companies

Measure	Technology Network	Company Network
Number of nodes	149	229
Number of edges	971	4067
Average degree	13.03	35.51
Number of connected components	7	4
Size of largest connected component	143	224
The average shortest path length	2.79	2.02

A subset of technologies from the larger network are used in this chapter:

- Car to car communication: commercialization expected in 2017, electronics
- Solar fuel: commercialization expected in 2017, started in 2007, biotechnologies

- Agile robots:

The other important metrics regarding individual nodes in the network are the centrality measures; degree, betweeness and closeness. For the purpose of this study we define the score of a node as the sum of these three measures and use it as the benefit of the node. The paths are characterized by the shortest lengths between two nodes and the shortest path represents the cost of moving from one node to another one.

We can also build a combined network of technologies and companies. As an example, Figure 5-2 shows the combined network and the shortest path between two technologies 'magic leap' and 'solar fuel' and the two companies which own these technologies respectively 'Magic leap' and 'Joule'. The network also shows the path between the source company 'Magic Leap' and the target technology 'solar fuel'.

The combined network shows alternative paths to reach a technology via a combination of technology nodes and company nodes. These can be seen as technology acquisitions and company partnerships . In fact we can see that the technology to technology path is the most costly (5.0) as these two technologies are in different domains while the company to company path is less costly (1.11) as companies can have diverse technology portfolios and may be connected to each other in other ways than the technologies considered. The least costly path (0.75) is one of mixed strategy using both technology and company links. This is more beneficial when the technology domains of the source and target nodes are different. The three respective paths are listed below:

- company to company path : ['Magic Leap', 'Rethink Robotics', 'CFM International', 'Joule']
- technology to technology path: ['magic leap', 'oculus rift', 'smart watches', 'personalized medical monitors', 'supergrids', 'solar fuel']
- company to technology path: ['Magic Leap', 'Rethink Robotics', 'CFM International', 'Snapchat', 'Cochlear', 'Flagship Ventures', 'solar fuel']

Next, we can look at a network where the links between companies are disabled while keeping the technology-to-technology links and the technology-to-company links. We note that the path using a mixed strategy is again the least costly but its cost has increased to 3.5 compared to the case where we have represented direct links between companies. The shortest path combining technologies and companies is shown on Figure 5-3 and includes the following nodes ['Magic Leap', 'magic leap', 'Google', 'project Loon', 'smart wind and solar', 'solar fuel'].

Firm positions in the network

Now that we have identified the important networks and metrics from **Chapter 4** we can use them to understand the position of a firm in the technology landscape.

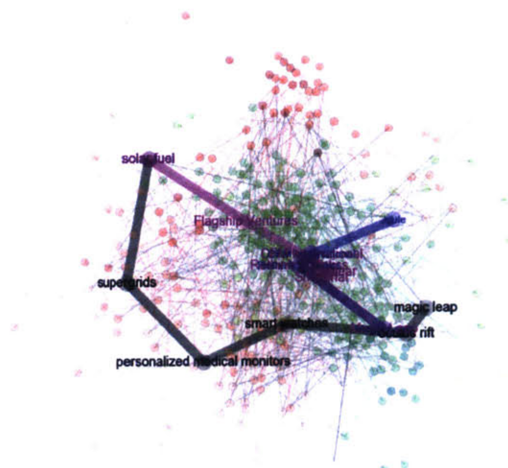

Figure 5-2: Paths between the Companies 'Magic Leap' and 'Joule' and between the Technologies 'magic leap' and 'solar fuel' Shown on a Combined Network of Technologies and Companies

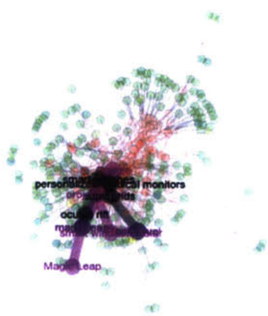

Figure 5-3: Paths between the Technologies 'magic leap' and 'solar fuel' and between the Company 'Magic Leap' and the technology 'solar fuel' Shown on a Combined Network of Technologies and Companies

In particular, we can identify the technologies related to a given firm and project them on the network. For instance, Figure 5-4 shows the nodes within Google's portfolio, Figure 5-5 shows the nodes within GM's portfolio and Figure 5-6 shows the nodes within both companies' portfolios. We note that the technology 'data mining' appears in GM's portfolio but is not present in Google's portfolio. One possible reason is the fact that data mining is not considered a breakthrough technology for a company like Google that is already doing it, whereas for GM this is a novel technology.

- Google portfolio: ['agile robots', 'apple pay', 'bayesian machine learning', 'cloud programming', 'Conversational Interfaces', 'deep learning', 'homomorphic encryption', 'Immune Engineering', 'internet dna', 'magic leap', 'mobile collaboration', 'neuromorphic chips', 'project Loon', 'real-time search', 'smart watches', 'social indexing', 'Tesla Autopilot']
- GM portfolio: portfolio ['car to car communication', 'data mining', 'solid state batteries', 'Tesla Autopilot']

Once the position of the firm in the network is known, we can see which of the nodes in its portfolio offers the best option to reach the target. In this case we take the same example of Google versus GM pursuing 'solar fuel'.

- Google's path: ['agile robots','robot design','synthetic biology','supergrids','solar fuel'] with a path length of 3.25.
- GM's path: ['car to car communication','mechatronics','synthetic biology','supergrids','solar fuel'] with a path length of 3.

Other examples of competitive positions include; (1) Google and Apple and (2) Google and Tesla competing fort car to car communication. There positions are shown in Figure 5-7 and Figure 5-8 respectively.

5.2 Strategic Games

In this section we will use simple examples of strategic moves to illustrate how the network is used to calculate the payoffs of different strategies. In particular, we will start with simple simultaneous games between two companies seeking the same technology. By varying the domains of the starting technology and the target technology we can compare how payoffs are impacted when a company plays within its technology domain or outside of it (new entrant versus incumbent).

Then we will explore games of commitment, where the firm uses a first mover's advantage to commit to a given strategy thereby altering the payoffs of the game. Commitment represents a strategic choice often faced in business and is particularly relevant to technology companies, as any credible commitment comes at a high investment cost.

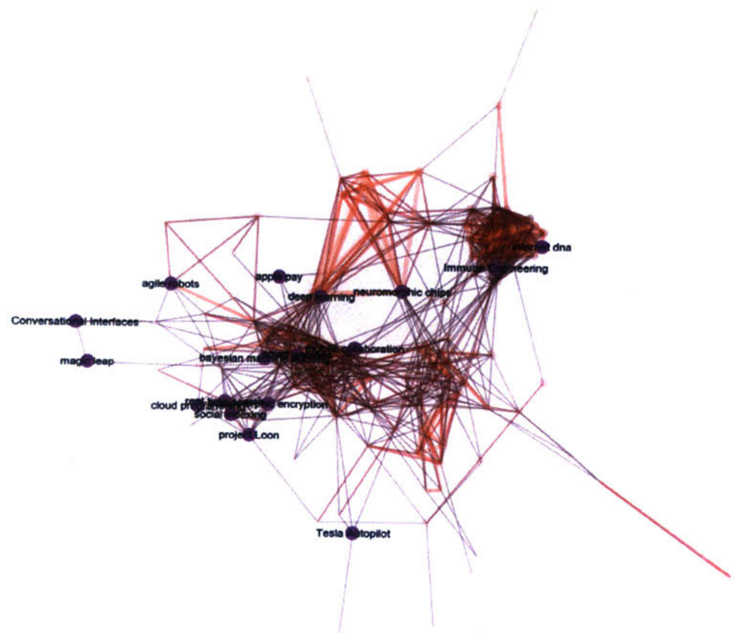

Figure 5-4: Nodes within Google's Network

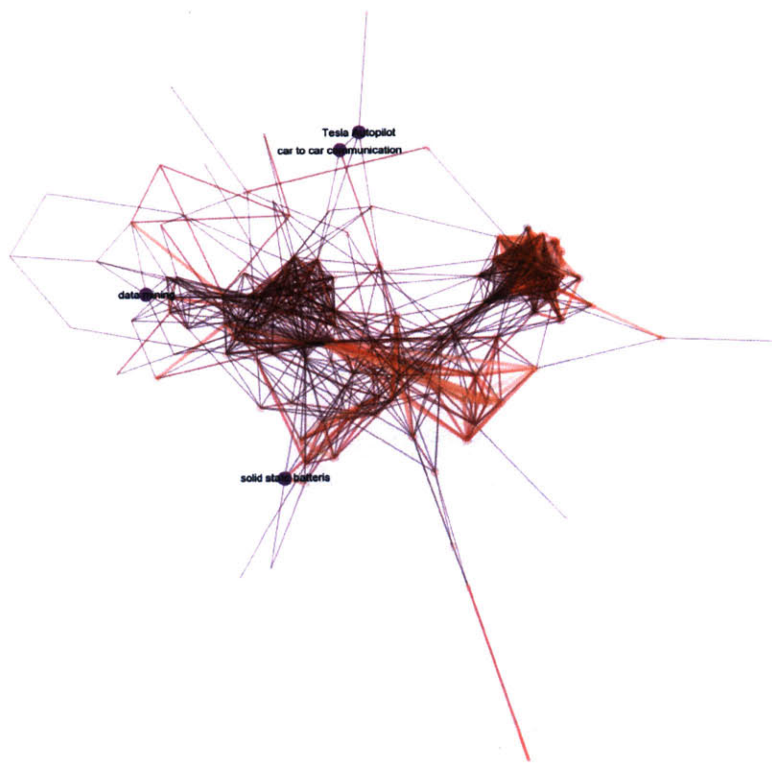

Figure 5-5: Nodes within GM's Network

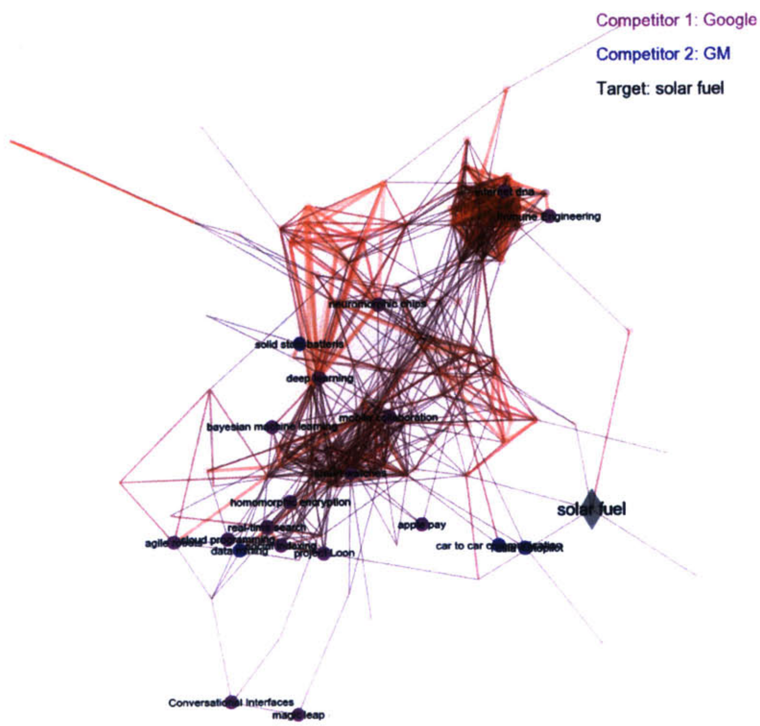

Figure 5-6: Nodes within Google's and GM's Networks

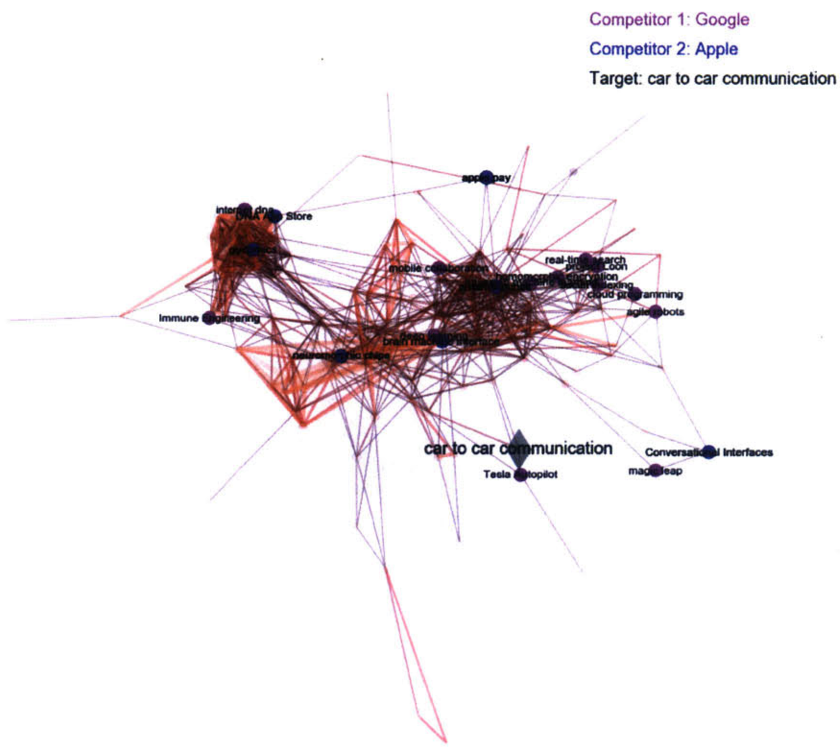

Figure 5-7: Google and Apple's Positions on the Technology Network

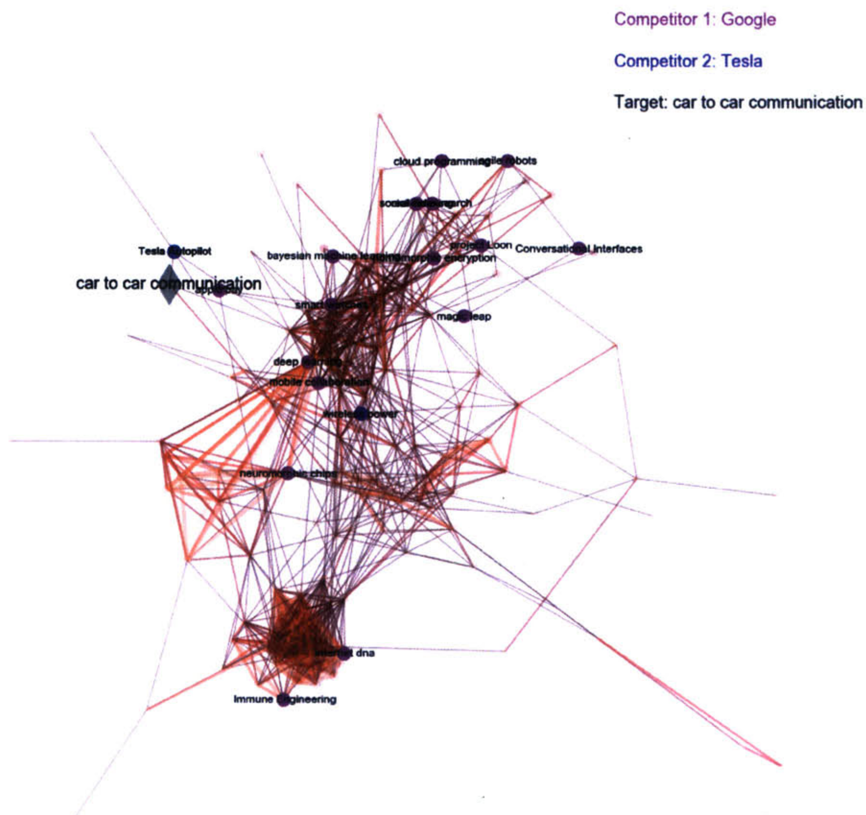

Figure 5-8: Google and Tesla's Positions on the Technology Network

We will also analyze games of cooperation where two companies can decide to collaborate and jointly develop a technology. This game is also called a "stag hunt", it is a two stage game in which the companies can collaborate or defect in the first stage and trust each other or not in the second stage.

Finally, since technology companies are often involved in acquisitions, we will examine a bargaining game with the sell off of a technology and two potential acquirers.

We will use game trees and payoff matrices to represent the moves and payoffs of the two players in the game. Once the game is set, the next step is to find the equilibrium that represents the best strategy each player can follow in response to its opponent's strategy. Dominant strategies represent the actions the player should take regardless of the opponent's actions. Backward induction will help us determine the best starting strategy.

Rules:

The players and technologies are respectively taken from the companies' and technologies' networks presented in the previous section. Depending on the game there will be two or three players. In games with two companies, each firm's starting position is a technology that belongs to its portfolio and the target technology is one that does not belong to either one of the competitors. In games with three players (limited to the price bargaining game), the seller determines the target technology and its price and the two competitors start from their respective technologies that are the "closest" to the target technology. Here, closest refers to the path length between the source and target technologies as defined in the network.

The payoffs of a game are calculated as follows:

- Given a target technology/product P0
- Given a company/Entity E1
- Given a company/Entity E2
- Find the list of products E1_Pn (n=1,N connected to Entity E1)
- Find the list of products E2_Pn (n=1,N' connected to Entity E2)
- Find the shortest path between P0 and any node in E1_Pn
- Find the shortest path between P0 and any node in E2_Pn
- Calculate the cost of the shortest path
- Calculate the benefit of the node
- Calculate the payoff for each company as (node benefit - path cost)
- With two or more companies investing the benefit of the node is divided by the number of known competitors.

5.2.1 Game 1 - SolarCity versus Siemens - Both in

In this hypothetical game we consider SolarCity and Siemens [46], two companies in the energy sector targeting a bio-fuel technology for acquisition. SolarCity is a start up offering solar energy services with a business model of leasing solar panels to homeowners and businesses. It was founded in July, 2006 and went public on the NASDAQ exchange in 2012 with a market capitalization of around $600 million. On the other hand, Siemens is an established multinational engineering and electronics company involved in the fields of industry, energy, transportation and healthcare. It reported a global revenue of approximately 73.5 billion euros for the year of 2011. Siemens went public in March, 2001.

Both companies are developing novel technologies to harness solar energy through higher efficiency solar panels. While SolarCity is betting on mass production of high efficiency solar panels based on existing technologies, Siemens is betting on a new technology that can improve solar panels' efficiency by more than two folds compared to current technologies.

The two companies have followed similar strategies of commitment; SolarCity by building a $750 million production plant located in Buffalo NY - the largest in North America - and Siemens by investing in R&D to develop advanced semiconductor materials with higher yields.

These investments confer the two companies leading positions in the sector of solar energy and to maintain their competitive advantage, it is expected that they will continue to evaluate novel technologies to expand their portfolios. We take solar fuel as an example of a novel technology these two companies could consider adding to their portfolios and analyze a game in which they compete against each other to to acquire it. Solar fuel is a technology that uses sunlight to efficiently convert carbon dioxide into ethanol or diesel [45].

The payoff matrix (Table 5.2 shows that both companies have a dominating strategy of investing as the payoffs are always positive compared to the case where the companies do not invest. This latter result is attributed to the fact that the cost for both companies is low as they make investments within their own technology domain. The game tree is shown in Figure 5-9.

Table 5.2: Payoff Table for SolarCity and Siemens Competing for the Technology Solar Fuel

SolarCity / Siemens	Invest	Do not invest
Invest	(1.16, 1.0)	(3.82, 0)
Do not invest	(0, 3.66)	(0, 0)

5.2.2 Game 2 - SolarCity versus Google - Both in

While Game 1 considered two competitors within the energy sector pursuing a bio-energy technology, in Game 2 we analyze the situation where one of the companies is from a different sector. Here, we define a game where SolarCity is still targeting the technology solar fuel but competing against Google instead of Siemens. Google is now part of Alphabet, a multinational corporation that is specialized in internet-related services and products. The company runs a semi-secret facility named Google X, where scientists

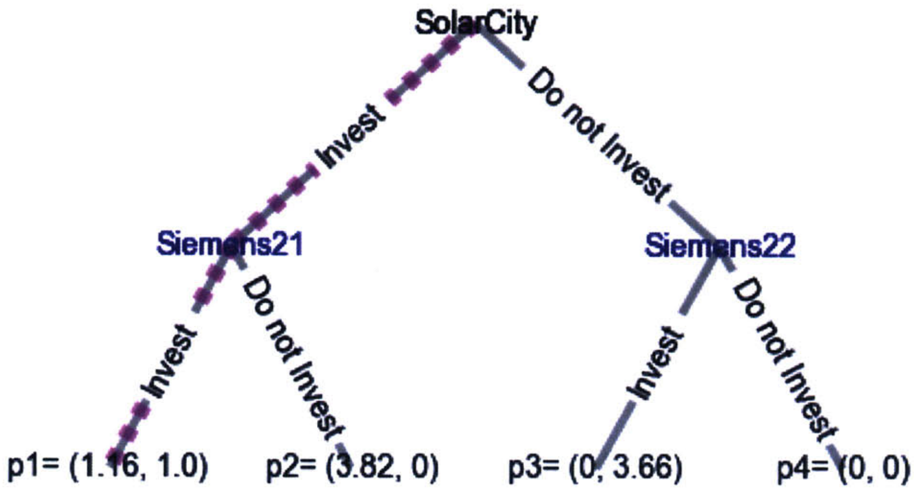

Figure 5-9: for SolarCity and Siemens Competing for the Technology Solar Fuel

and entrepreneurs work on improving breakthrough technologies by a factor of ten. Their projects include; Project Glass, the Google driveless car, Project Loon, and Google Contact Lens [46]. This layered corporate structure, with different strategic goals, has led Google to pursue investments in multiple technology domains outside its core sector of Information technology.

Solar fuel, the target technology is at the intersection of energy and bio-technologies, two sectors in which Google has invested in the past. In this particular game, the recommended path for Google to attain the target technology starts from internet DNA, a technology within Google's portfolio. Internet DNA is a system for trading genetic information it helps automate the comparison of DNA and thereby offers access to detailed genome information. This technical capability aligns with the underlying science behind solar fuel which is based on manipulating genes to create photosynthetic microorganisms capable of producing bio-fuels.

The payoff matrix (Table 5.3) in this case shows a dominating strategy for both SolarCity and Google to invest as the payoffs are positive independently of what the opponent does. This is not always the case when one of the competitors is from a different industry sector, but Google has a portfolio that spans the entire network of technologies and allows the firm to enter new sectors. This rather beneficial competitive position followed prior strategic investment to diversify Google's portfolio. Figure 5-10 shows the game tree.

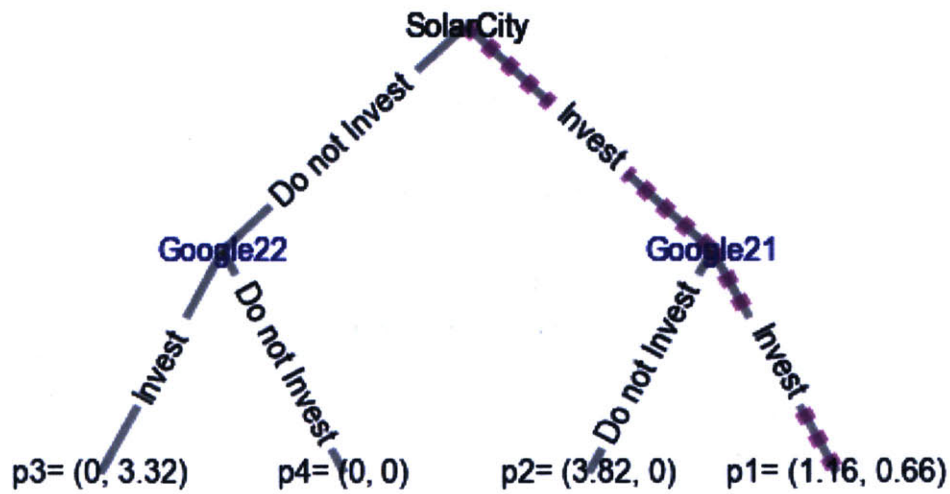

Figure 5-10: for SolarCity and Google Competing for the Technology Solar Fuel

Table 5.3: Payoff Table for SolarCity and Google Competing for the Technology Solar Fuel

SolarCity / Google	Invest	Do not invest
Invest	(1.16, 0.66)	(3.82, 0)
Do not invest	(0, 3.32)	(0, 0)

5.2.3 Game 3 - Google versus Magic Leap - Google In, Magic Leap Out

In Game 3 we consider two companies in the information technology sector; Google and Magic Leap competing to invest in the same a bio-energy technology, solar fuel. As we saw in Game 2, Google has a dominating strategy to invest in this technology if facing a single competitor. This is was possible due to Google's diverse portfolio that allowed it to enter different technology domains than its core sector while keeping the payoffs always positive.

In Game 3, we will illustrate that a company with a small and specialized portfolio in Information technology does not enjoy the same strategic competitive position when entering a different sector. As an example we consider Magic Leap which commercializes a proprietary virtual reality wearable technology.

According to the payoffs shown in Table 5.4 Google has a dominating strategy of investing regardless of the opponent's strategy resulting in Magic Leap's best strategy to not invest. We also note that even when investing alone the payoff is low. This result illustrates how companies with specialized, small portfolios have a weaker competitive position when entering different sectors than those with diverse portfolios.

Table 5.4: Payoff Table for Magic Leap and Google Competing for the Technology Solar Fuel

Google / Magic Leap	Invest	Do not invest
Invest	(0.66, -2.34)	(3.32, 0)
Do not invest	(0, 0.32)	(0, 0)

5.2.4 Game 4 -Apple versus Toyota - A Game of Chickens

As the trend of self-driving cars continues to increase, many companies are working on the technologies that will enable autonomous vehicles. In this race to innovate, traditional car manufacturers and in particular Toyota have taken the lead as shown in Figure 5-12. Apple, on the other hand with just one invention across all areas of self-driving innovation, in telematics, is trailing the race despite all the headlines. This is not surprising as Apple's core sector is the design and manufacturing of electronics and the development of custom software. It has a dominant market presence in the mobile communication, media devices and personal computers. but only recently started building a driveless car [46]. Toyota on the other hand is an automobile manufacturer operating 52 overseas manufacturing subsidiaries in 27 countries and regions and is the overall global leader in autonomous automotive innovation.

While driving assistance features already exist in many cars sold today, the leap to fully autonomous vehicles is yet to happen. Autonomous driving involves not only propelling a vehicle winch is where the expertise of traditional car manufacturers such as Toyota lies. But it also involves navigating a vehicle

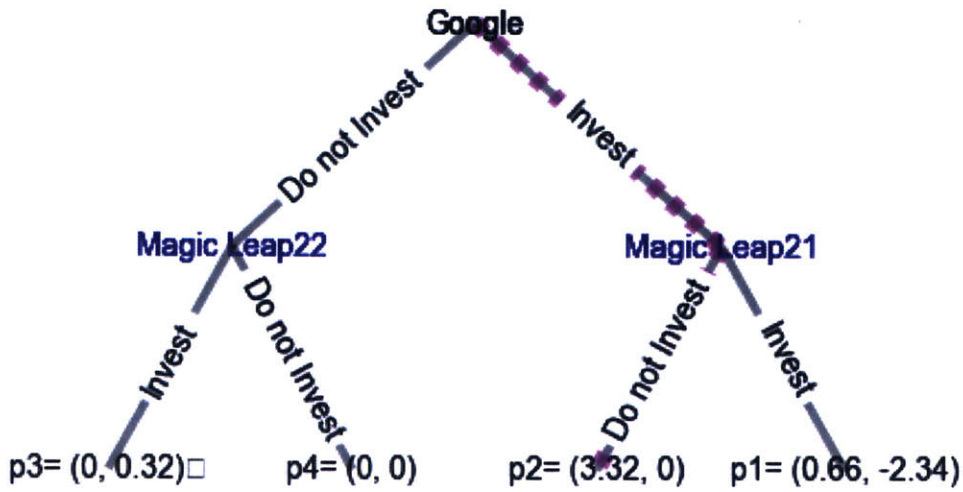

Figure 5-11: for SolarCity and Google Competing for the Technology Solar Fuel

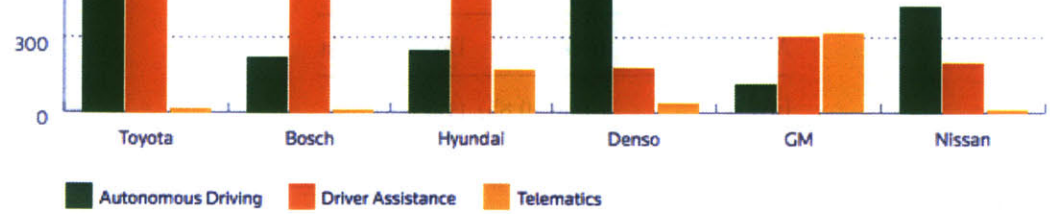

Figure 5-12: Self-driving Cars top Innovators 2010 to 2015

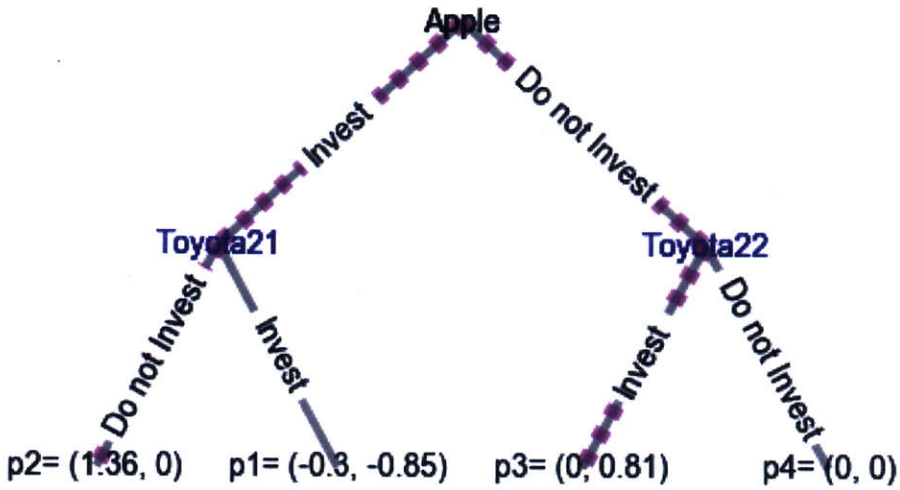

Figure 5-13: Game Tree for Apple and Toyota Competing for the Technology car to car communication

without human input through the use of sensors and control instruments. These controls rely heavily on custom software , which is where Information Technology companies such as Apple have an advantage.

In this game, Apple starts from a wireless technology that enables human machine interactions while Toyota starts from advanced batteries as both target the technology known as car to car communication.

The payoff table (Table 5.12) indicates that this is a coordination game. If Apple invests, Toyota should not invest and vice versa. Therefore the actions and outcomes of one player depend entirely on what the opponent will do, there are two nash-equilibria. The game tree is shown in Figure 5-13.

Table 5.5: Payoff Table for Apple and Toyota Competing for the Technology car to car communication

Apple / Toyota	Invest	Do not invest
Invest	(-0.3, -0.85)	(1.36, 0)
Do not invest	(0, 0.81)	(0, 0)

Since there is incomplete information as none of the opponents knows what the other will do, we can look at the payoff of one as a function of the probability of the other investing.

In Figures 5-14 and 5-15 we can see that Apple should invest as long as the probability that Toyota will

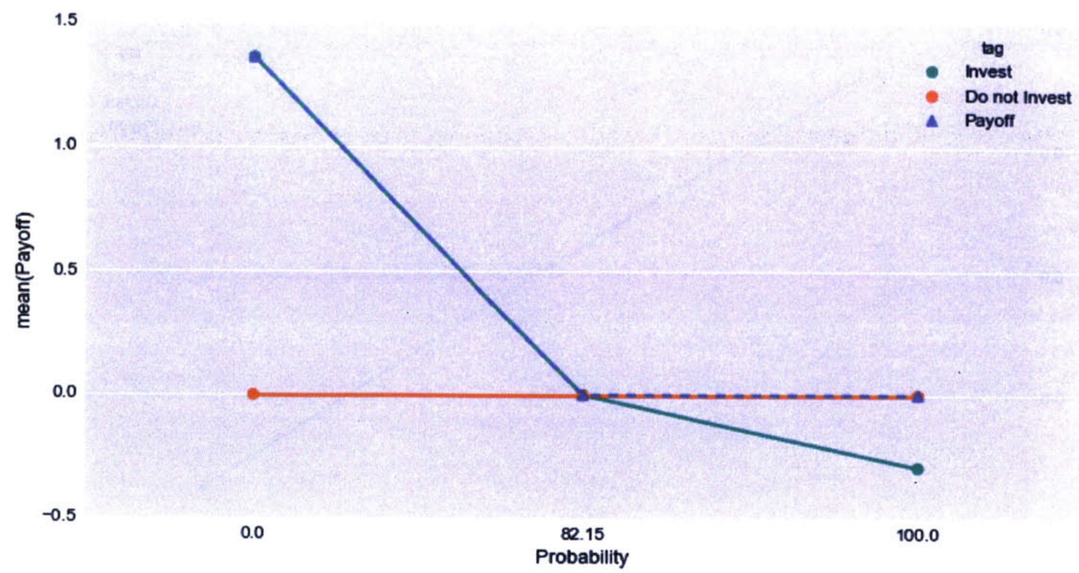

Figure 5-14: Apple's Investment Strategy

invest remains below 82%, while Toyota should invest as long as the probability that Apple invests remains below 49%. With better knowledge of the willingness to pay and the consideration of other strategic decision, stakeholders within the two companies can determine which outcome is more likely and act accordingly.

5.2.5 Game 5 - Apple and Tesla - a Game of Cooperation

Game 5 illustrates a scenario described by analysts in a recent report on the state of self-driving automotive innovation [47]. According to this report , to enter the self-driving market, Apple needs to partner with an automotive manufacturer. In particular, analysts predicted that a strong partner would be Tesla, according to a review of both company's patent portfolios.Therefore, in this game, we set to explore a strategic game in which the two companies can "collaborate" or "defect" in a game theoretic sense.

Tesla Motors is an electric car company started by Elon Musk whose vision is to disrupt the electric vehicle market and serve at a catalyst for the market. Tesla is poised to play a major role in the autonomous vehicle market as its latest software updated, dubbed Tesla's autopilot, delivered to its fleet of approximately 60,000 cars enabled the vehicles to drive autonomously. The software update enabled the embedded sensors to gather real-time data and use it to manage speed, steering to change lanes, and park itself. While most of these features are present in other competitors' cars, the self-steering is the true advance toward driving autonomy.

The payoff table (Table 5.6) shows that Tesla has a dominating strategy of investing resulting in Apple not investing. The game tree is shown in Figure 5-16.

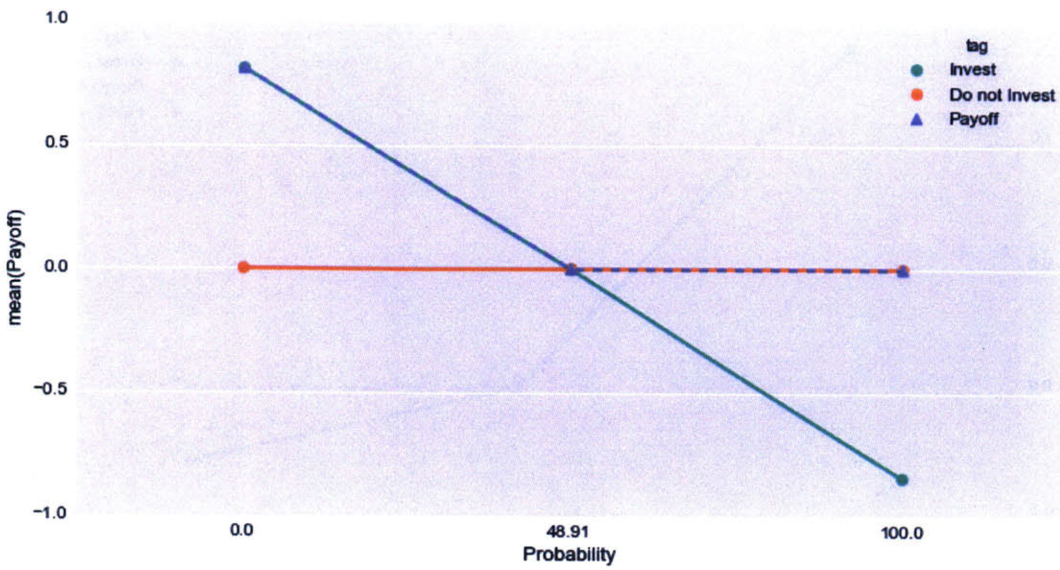

Figure 5-15: Toyota's Investment Strategy

Table 5.6: Payoff Table for Apple and Tesla Competing for the Technology car to car communication

Apple / Tesla	Invest	Do not invest
Invest	(-0.3, 0.65)	(1.36, 0)
Do not invest	(0, 2.31)	(0, 0)

Now, considering the case with collaboration, and to better understand why a collaboration between the two companies would be beneficial for both, we have modified the payoffs in the following way:

- Collaboration: payoffs(Apple)=payoffs(Tesla)= benefit/2-cost/2
- Both defect: payoffs(Apple)=payoffs(Tesla)= benefit/2-cost
- One defects: payoffs(defector)= benefit-cost (as if invested alone) and payoffs(collaborator)= 0

The payoffs listed above are based on two assumptions; first when collaborating both firms capture half of the benefit but reduce their cost by half as they can use the complementary technologies the other one has as opposed to developing everything on their own. Second, when both defect, we are back to the competitive game payoffs. Finally when only one defects, we are back at the scenario where only one of the players invests.

As we can see in Table 5.7, The outcome of both collaborating is better than the outcome of both defecting, but the outcome of each one of them is better when defecting than collaborating. In repeated games, the outcome of this first stage of cooperation will determine whether the two companies trust each other.

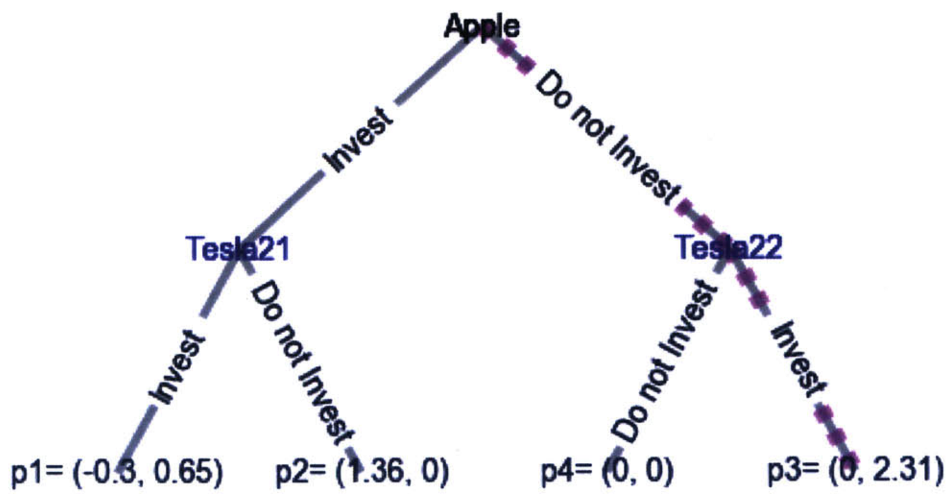

Figure 5-16: Game Tree for Apple and Tesla Competing for the Technology car to car communication

Table 5.7: Game Tree for Apple and Tesla Collaborating for the Technology "car to car communication"

Apple / Tesla	Collaborate	Do not collaborate
Collaborate	(0.68, 1.15)	(0, 2.31)
Do not collaborate	(1.36, 0)	(-0.3, 0.65)

5.2.6 Game 6 Amazon and Toyota - The value of Commitment and First Mover's Advantage

On March 21, 2016, Bloomberg reported that Google was selling off one of the robotics companies it acquired in 2014 and presumably the one with the most promising technology. Google acquired Boston Dynamics in 2014 when it went on a spree to acquire key technologies to make more advanced robots. Analysts speculated at the time that the company's intent was to build advanced new robots for factories and homes.

Boston Dynamics developed the 'agile robots' technology, the two-legged humanoid and four-legged cheetah robots were developed with support from DARPA, initially for military applications. Boston Dynamics began as a spin-off from the Massachusetts Institute of Technology, the first robots developed were capable of running and maneuvering like animals. The company was founded in 1992, and continues to push the limits of ground-breaking robotics. However, product development and commercialization prospects remained slow and many analysts see these issues as the main reason for Google's divestment.

The Toyota Research Institute and Amazon are possible acquirers. Toyota's interest in robotics aligns with the automation needs in the car manufacturing sector although the company may be price sensitive after investing close to $1Billion in a new Research and Development center. Amazon's interest in robots lies with the automation needs of its fulfillment centers. Amazon started as an e-commerce retailer providing consumers merchandise and content from resellers and third-parties. In March 2012, Amazon acquired Kiva Systems for 775 million dollar, a manufacturer of robots for fulfillment centers started in 2003. Since then Amazon focused on internal development and deployment of the technology.

Amazon and Toyota's technology portfolios are shown in Figure 5-17 as well as the target technology developed by Boston Dynamics 'agile robots.

The respective paths for reaching the technology 'agile robots' for both companies are shown in the Figures 5-18 and 5-19. Starting with Toyota, we can identify three types of connections to Boston Dynamics and the technology 'agile robots'. One path on the technology network, one path on the company network and one path that is hybrid going through both.

In this game we will analyze a situation in which Toyota has a first mover's advantage to make an offer to Google to acquire the target technology and Amazon can either stay out or invest and compete.

- Google: set the selling price for the target technology: high or low

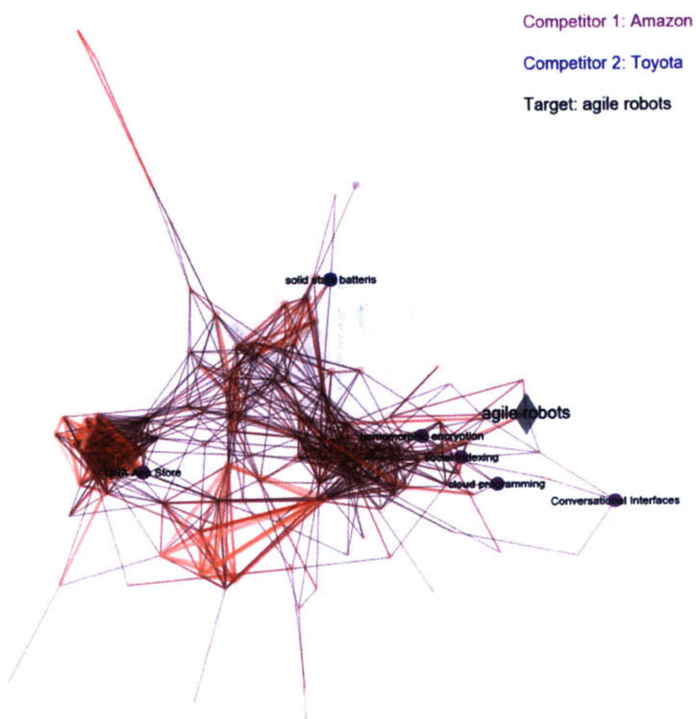

Figure 5-17: Amazon and Toyota Positions on the Technology Network

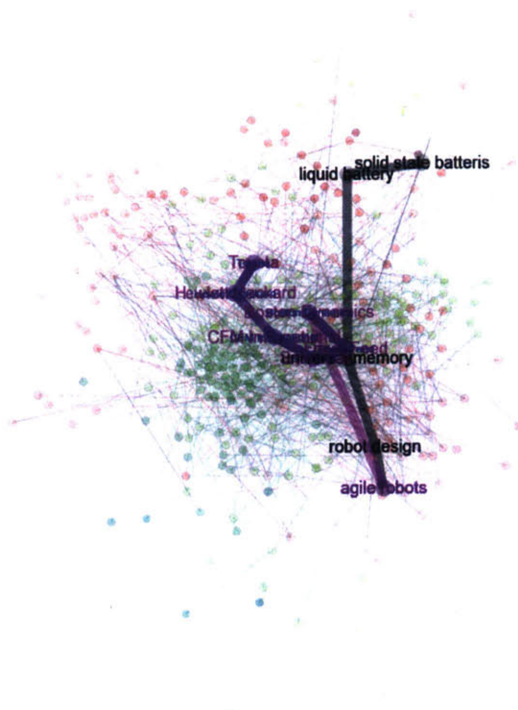

Figure 5-18: Toyota's path to "agile robots" on the Technology Network

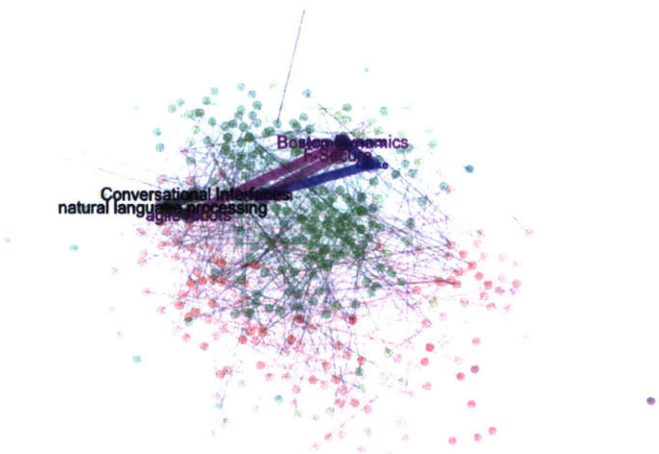

Figure 5-19: Amazon's path to "agile robots" on the Technology Network

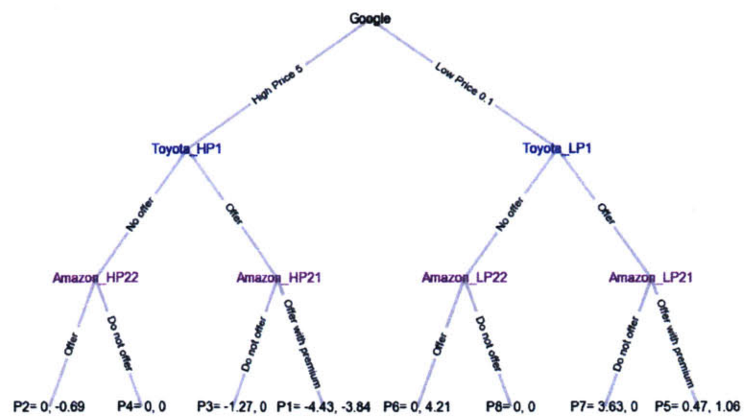

Figure 5-20: Game Tree Representing Google, Toyota and Amazon's Actions in the sell-of of the Technology "agile robots"

- Toyota has the first mover's advantage to make an offer or not: offer or no offer
- Amazon has the option to stay out, match Toyota's offer or offer a premium: stay out, match, outbid

First, we consider the decision Google has to make on setting the price to a low, medium or high value. At the extreme high price, the outcomes for both companies are either negative if they invest or zero if they do not. Therefore, neither company will invest and Google does not sell the technology. On the other side of the spectrum, if the price is very low, both companies Amazon and Toyota would stay in the race and both would invest as the payoffs are always positive compared to zero when not investing. The game tree is shown in Figure 5-19.

The cases of interest are those where the price is in between these two extreme values. To better understand how the strategies of the two competitors change as functions of the price, we can look at a few examples.

Scenario 1 with a low price (0), both companies have a dominant strategy to invest.

Table 5.8: Payoff Matrix for the acquisition of "agile robots" at a price of 0

Toyota / Amazon	Invest	Do not invest
Invest	(0.57, 1.16)	(3.73, 0)
Do not invest	(0, 4.31)	(0, 0)

Scenario 2 with a low price (1), Amazon has a dominant strategy to invest and in that case Toyota should not invest.

Table 5.9: Payoff Matrix for the acquisition of "agile robots" at a price of 1

Toyota / Amazon	Invest	Do not invest
Invest	(-0.43, 0.16)	(2.73, 0)
Do not invest	(0, 3.31)	(0, 0)

Scenario 3 with a medium price (2), it is a game of chicken where one should not invest if the other invests.

Table 5.10: Payoff Matrix for the acquisition of "agile robots" at a price of 2

Toyota / Amazon	Invest	Do not invest
Invest	(-1.43, -0.84)	(1.73, 0)
Do not invest	(0, 2.31)	(0, 0)

Scenario 4 with a high price (3), Toyota has a dominant strategy to not invest, which in turn means that Amazon should invest.

Table 5.11: Payoff Matrix for the acquisition of "agile robots" at a price of 3

Toyota / Amazon	Invest	Do not invest
Invest	(-2.43, -1.84)	(0.73, 0)
Do not invest	(0, 1.31)	(0, 0)

Scenario 5 with a high price (4), Neither Toyota nor Amazon should invest as both have a dominant strategy to not invest.

Table 5.12: Payoff Matrix for the acquisition of "agile robots" at a price of 4

Toyota / Amazon	Invest	Do not invest
Invest	(-3.43, -2.84)	(-0.27, 0)
Do not invest	(0, 0.31)	(0, 0)

Scenario 5 with a high price (5), Neither Toyota nor Amazon should invest as both have a dominant strategy to not invest.

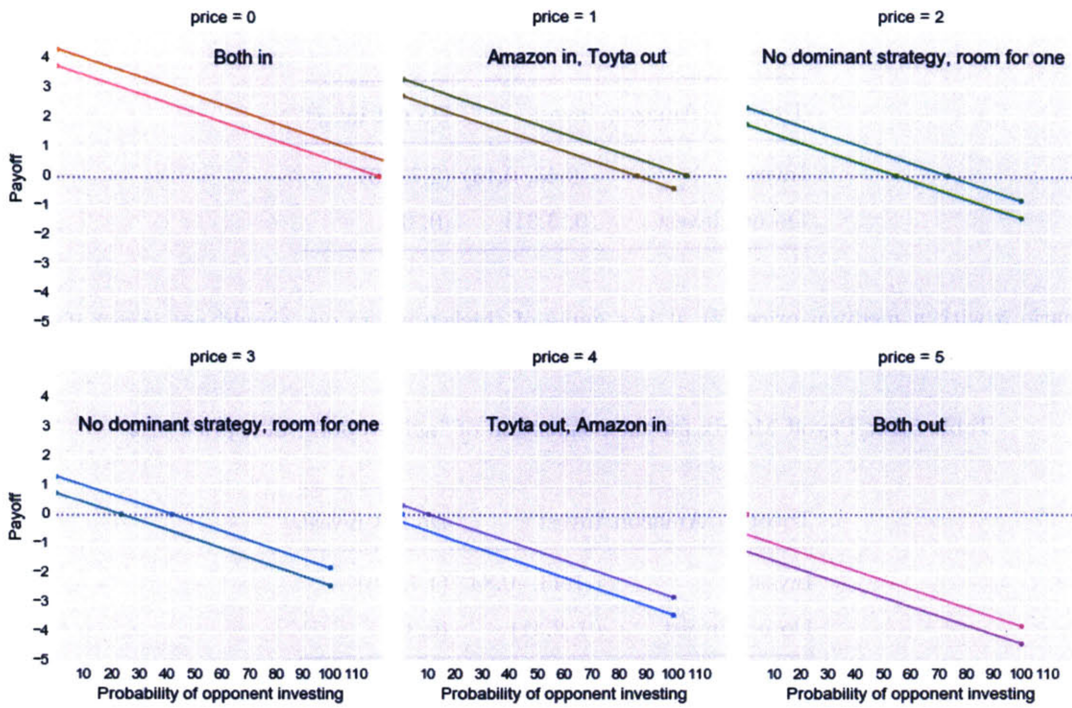

Figure 5-21: Strategies for the acquisition of "agile robots" at Different Price Points

Table 5.13: Payoff Matrix for the acquisition of "agile robots" at a price of 5

Toyota / Amazon	Invest	Do not invest
Invest	(-4.43, -3.84)	(-1.27, 0)
Do not invest	(0, -0.69)	(0, 0)

Figure 5-21 below generalize how the strategies for both players change under changing prices.

When both companies have dominant strategies, they will act accordingly. When the price range is approximately between 1 and 4, there is no dominant strategy and Toyota can play its first mover's advantage to alter the payoffs. By making a commitment in the form of an investment, Toyota would irreversibly alter the payoffs of the game leaving no option but to carry out the action of investing in the target technology in its self-interest. For instance, assuming a starting case where the price is 3 and no dominant strategy for either one of the competitors. One form of commitment is for Toyota to invest in technologies that can move it closer to the target technology. In this case we use the path in the technology network from the source technology to the target technology['solid state batteries','liquid battery','universal memory','robot design','agile robots']. We assume that Toyota would make a commitment by investing in the first two technologies along this path to move itself closer to the target technology.

This move has a cost equal to the path length between the initial position and the new potion ['solid state batteries','liquid battery','universal memory']. As we can see from the payoffs table (Table 5.14), the new dominant strategy for Toyota is to invest, thereby forcing amazon to not invest.

Table 5.14: Payoff Matrix for the acquisition of "agile robots" at a price of 3 with Prior Commitment from Toyota

Toyota/Amazon	Invest	Do not invest
Invest	(-1.09, -1.84)	(2.06, 0)
Do not invest	(-1.33, 1.31)	(-1.33, 0)

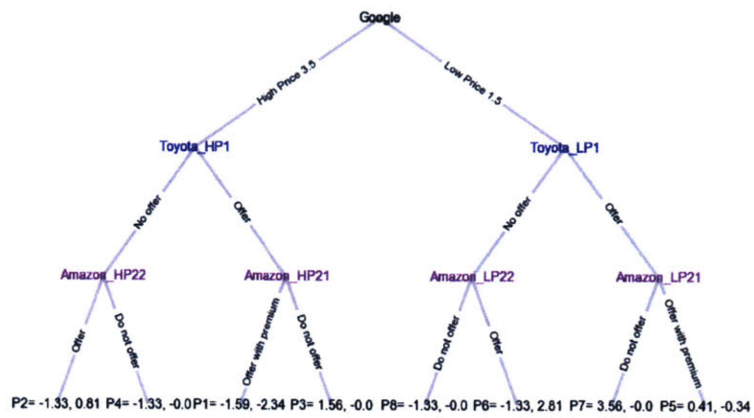

Figure 5-22: Game Tree Representing Google, Toyota and Amazon's Actions in the sell-of of the Technology "agile robots - With Prior Commitment from Toyota"

The new game tree shows that for a starting price without dominant strategies for both competitors, by moving first and making a commitment, Toyota ensures that it is the only investor in the targeted technology.

Chapter 6

Conclusion

Technology companies are continuously screening for, or developing, breakthrough technologies. The resulting innovative products and services often are developed by tapping and aligning internal know-how with external technologies. In this research, we sought to improve the process for selecting external breakthrough technologies by understanding the similarities between these technologies using network theory. Additionally, employing game theory to improve strategic decisions, we applied this newly-acquired knowledge about technology networks to provide competitive advantage to technology firms.

First, we developed a systematic analysis framework that analyzes multiple dimensions of similarities to create networks of breakthrough technologies. As a firm considers a particular target technology, the framework offers a method for calculating the payoff of following a particular path in the network to attain the target from any technologies already in the firm's portfolio. In competitive scenarios, the framework is used to analyze competitors' dominant strategies and likely moves.

We compared various publications to extract lists of breakthrough technologies; then selected the *MIT Technology Review* as a source of textual descriptions of these technologies. Next, using an API to the IBM Watson suite, we conducted a semantic analysis of these technologies to extract four categories of features; concepts, keywords, taxonomy and entities, for each technology considered.

We employed these four categories of features as dimensions of similarities between the technologies, and compared the characteristics of the resulting networks. In these networks, breakthrough technologies are represented by nodes, while dimensions of similarities are represented by links between these nodes. 'Concepts' were found to be the most relevant dimension of similarity, so they were selected to represent technology networks throughout the remainder of the research. While natural language processing API returns the relevance of each extracted feature, these relevance scores were not incorporated in the calculation of the weight of a link. Instead, the weight of a link between two technologies is proportional to the number of concepts in common. These dimensions of similarities were particularly helpful in categorizing the technologies in a five by five matrix of elementary processes and operands. Further, the concepts network revealed a clustered structure of technology communities. When analyzed over time, these communities showed they initially evolved separately, building links within individual communities, before branching out to other technology domains.

Similarly, we built networks of companies along two dimensions; concepts and technologies. These networks allowed us to uncover underlying connections between companies, and in some cases, offered shorter paths to connecting two technologies.

We also introduced two metrics; (1) the benefit of a node, derived based on the node's position in the network, and (2) the cost of a link, estimated based on similarities of technologies it connects. The benefit of the node was simply calculated as the sum of its degree, betweenness and closeness, but other representations of a node's value were not considered.

The model is used to derive a firm's position in the network, and its target technology. It recommends which technology in the firm's portfolio should be used to start the firm's walk in the network, and recommends the shortest path. When considering competitive games, the model provides the payoff matrix of both competitors, accounting for competitive games versus collaborative games. Finally, the application of this method is illustrated with six use cases that analyze the effect of technology domains on the payoffs, and the types of strategies used by the competitors.

Chapter 7

Future Work

In this analysis we developed a conceptual prototype of a framework to support technology strategy decisions. In this first stage, we developed a simple model which does not include the notion of technology risk, discounted cash flow for technology investments and uncertainty on technology benefit. Further, the model uses a limited set of inputs in the form of breakthrough technologies and the companies linked to them. Finally, the model was used with a limited set of strategic games, some of which were hypothetical and limited the set of strategic decisions to; investing alone, collaborating with a competitor and committing to a particular technology through early investment. In this section we discuss the limits of the modeling decisions made and how they can be further expanded and validated.

Technology Infusion In this analysis we assume that the benefit of a node is captured when it is reached through the recommended path. However, as discussed in **Chapter 2** , the benefit of a technology is only captured if it is successfully integrated with other systems to deliver the intended service or product. While this analysis does not explicitly account for the integration step in the benefit calculation, it assumes that if a company follows the path to a given target technology it will build the needed technical capability and know-how along the way to be able to integrate it successfully. In contrast, if a company acquires a technology it has no direct or indirect connections to, it is an indication that the existing portfolio of the company does not share technical similarities with the target technology. The inclusion of a quantitative measure of technology infusion will be an important step in validating this framework.

Technology Position To better represent companies' positions in the network, the actual portfolio of the firm needs to be represented. In the current model, the position of the firm only includes those technologies within the used dataset that are linked to the company.

Technology Risk Technology Readiness Levels (TRL) were not considered in the calculations of benefit and costs in this analysis. TRL(s) can introduce high levels of risks and play a major role in the technology strategy of a firm. Low TRL(s) demand higher level of sustained investment to bring the technology to maturity, clear regulatory requirements if they exist and demonstrated a working proof of concept. In addition, technologies with low TRL(s) may prove more challenging to integrate as they may use novel processes, require new interfaces or call for new capabilities. The dataset of technologies used in this analysis originates from a single source, the "MIT Technology Review", and to be included in this list, a given technology should have demonstrated some level of feasibility and achieved a successful milestone. However, TRL(s) should be added and used as a systematic tool to compare technology maturity. One way the technology risk can be accounted for would be to penalize the paths that cross nodes with low TRL(s) and favor those that cross nodes with high TRL(s).

Technology Benefit In this analysis, technology benefit is based solely on network metrics and in particular nodes' centrality measures. This is in contrast to Net Present Value and Options calculations that are more traditionally used. The concept of benefit defined here is meant to represent the intangible benefit a given technology confer by being versatile in its use and allowing the firm to hedge risk by using the technology in different domains. To illustrate this point, one can compare edge nodes with central nodes. Edge nodes represent technologies that are very specialized and cannot readily be applied to other domains. In contrast central nodes are ones that share sufficient elementary functions with other technologies thereby enabling the firm to easily move into an adjacent technology domain.

However, this definition of technology benefit by itself is not sufficient to inform the final stages of technology selection and thus needs to be combined with NPV and options calculations. One such way would be to include the price of the technology calculated as an option coming to maturity at the expected time of commercialization.

Technology Path Cost In the path cost calculation, the weight of a path between two technologies represents the number of concepts in common. One additional parameter that can be added, is the relevance of these concepts to each technology. This additional parameter allows the links to be more relevant.

Payoff Calculation The current payoff calculation is based on deterministic benefit and cost, however as previously discussed, the notions of market and technology risks introduce uncertainty in the benefits and costs. This probabilistic calculation needs to be added to the model for more realistic results.

Strategic Games Finally, in the strategic games considered, the options were limited to coordination, collaboration and commitment. But there are many other strategic decisions available to the firm that are not represented in the games proposed. These include; technology licensing, corporate equity investment, IP "trolling", alternative technologies and so forth. In addition, in commitment games, this analysis only considered one competitor's actions and did not model a case with sequential moves where both competitors use signaling and commitment. Such a game is better represented by a Markov Perfect Equilibrium model.

Appendix A

Tables

Table A.1: Technologies extracted from the MIT Technology Review - Full List

	Product	year	Link
0	Immune Engineering	2016	https://www.technologyreview.com/s/600763/10-b...
1	Precise Gene Editing in Plants	2016	https://www.technologyreview.com/s/600765/10-b...
2	Conversational Interfaces	2016	https://www.technologyreview.com/s/600766/10-b...
3	Reusable Rockets	2016	https://www.technologyreview.com/s/600767/10-b...
4	Robots That Teach Each Other	2016	https://www.technologyreview.com/s/600768/10-b...
5	DNA App Store	2016	https://www.technologyreview.com/s/600769/10-b...
6	SolarCity Gigafactory	2016	https://www.technologyreview.com/s/600770/10-b...
7	Slack	2016	https://www.technologyreview.com/s/600771/10-b...
8	Tesla Autopilot	2016	https://www.technologyreview.com/s/600772/10-b...
9	Power from the Air	2016	https://www.technologyreview.com/s/600773/10-b...
10	apple pay	2015	https://www.technologyreview.com/s/535001/appl...
11	brain organoids	2015	https://www.technologyreview.com/s/535006/brai...
12	car to car communication	2015	https://www.technologyreview.com/s/534981/car-...
13	internet dna	2015	https://www.technologyreview.com/s/535016/inte...
14	liquid biopsy	2015	https://www.technologyreview.com/s/534991/liqu...
15	magic leap	2015	https://www.technologyreview.com/s/534971/magi...
16	megascale desalination	2015	https://www.technologyreview.com/s/534996/mega...
17	nano-architecture	2015	https://www.technologyreview.com/s/534976/nano...
18	project Loon	2015	https://www.technologyreview.com/s/534986/proj...
19	supercharged photosynthesis	2015	https://www.technologyreview.com/s/535011/supe...

20	agile robots	2014	https://www.technologyreview.com/s/526536/agil...
21	agricultural drones	2014	https://www.technologyreview.com/s/526491/agri...
22	brain mapping	2014	https://www.technologyreview.com/s/526501/brai...
23	genome editing	2014	https://www.technologyreview.com/s/526511/geno...
24	microscale 3-D printing	2014	https://www.technologyreview.com/s/526521/micr...
25	mobile collaboration	2014	https://www.technologyreview.com/s/526526/mobi...
26	neuromorphic chips	2014	https://www.technologyreview.com/s/526506/neur...
27	oculus rift	2014	https://www.technologyreview.com/s/526531/ocul...
28	smart wind and solar	2014	https://www.technologyreview.com/s/526541/smar...
29	ultraprivate smartphones	2014	https://www.technologyreview.com/s/526496/ultr...
30	additive manufacturing	2013	https://www.technologyreview.com/s/513716/addi...
31	baxter the blue collar robot	2013	https://www.technologyreview.com/s/513746/baxt...
32	big data from cheap phones	2013	https://www.technologyreview.com/s/513721/big-...
33	deep learning	2013	https://www.technologyreview.com/s/513696/deep...
34	memory implants	2013	https://www.technologyreview.com/s/513681/memo...
35	prenatal DNA sequencing	2013	https://www.technologyreview.com/s/513691/pren...
36	smart watches	2013	https://www.technologyreview.com/s/513376/smar...
37	supergrids	2013	https://www.technologyreview.com/s/513736/supe...
38	temporary social media	2013	https://www.technologyreview.com/s/513731/temp...
39	ultra-efficient solar power 2	2013	https://www.technologyreview.com/s/513671/ultr...
40	3-D transistors	2012	http://www2.technologyreview.com/article/42767...
41	a faster Fourier Transform	2012	http://www2.technologyreview.com/article/42767...
42	crowdfunding	2012	http://www2.technologyreview.com/article/42767...
43	egg stem cells	2012	http://www2.technologyreview.com/article/42767...
44	facebook timeline	2012	http://www2.technologyreview.com/article/42767...
45	high speed material discovery	2012	http://www2.technologyreview.com/article/42767...
46	light field photography	2012	http://www2.technologyreview.com/article/42767...
47	nanopore sequencing	2012	http://www2.technologyreview.com/article/42767...
48	solar micro-grids	2012	http://www2.technologyreview.com/article/42767...
49	ultra-efficient solar power 1	2012	http://www2.technologyreview.com/article/42767...
50	cancer genomics	2011	http://www2.technologyreview.com/article/42368...
51	cloud streaming	2011	http://www2.technologyreview.com/article/42368...
52	crash-proof code	2011	http://www2.technologyreview.com/article/42369...
53	gestural interfaces	2011	http://www2.technologyreview.com/article/42368...
54	homomorphic encryption	2011	http://www2.technologyreview.com/article/42368...

55	separating chromosomes	2011	http://www2.technologyreview.com/article/42369...
56	smart transformers	2011	http://www2.technologyreview.com/article/42368...
57	social indexing	2011	http://www2.technologyreview.com/article/42368...
58	solid state batteris	2011	http://www2.technologyreview.com/article/42368...
59	synthetic cells	2011	http://www2.technologyreview.com/article/42369...
60	cloud programming	2010	http://www2.technologyreview.com/article/41854...
61	dual-action antibodies	2010	http://www2.technologyreview.com/article/41854...
62	engineered stem cells	2010	http://www2.technologyreview.com/article/41853...
63	green concrete	2010	http://www2.technologyreview.com/article/41854...
64	implantable electronics	2010	http://www2.technologyreview.com/article/41854...
65	light-trapping photovoltaics	2010	http://www2.technologyreview.com/article/41854...
66	mobile 3-D	2010	http://www2.technologyreview.com/article/41853...
67	real-time search	2010	http://www2.technologyreview.com/article/41853...
68	social tv	2010	http://www2.technologyreview.com/article/41854...
69	solar fuel	2010	http://www2.technologyreview.com/article/41853...
70	$100 genome	2009	http://www2.technologyreview.com/article/41218...
71	biological machines	2009	http://www2.technologyreview.com/article/41218...
72	hashCache	2009	http://www2.technologyreview.com/article/41219...
73	intelligent software assistant	2009	http://www2.technologyreview.com/article/41219...
74	liquid battery	2009	http://www2.technologyreview.com/article/41219...
75	nanopiezoelectronics	2009	http://www2.technologyreview.com/article/41219...
76	paper diagnostics	2009	http://www2.technologyreview.com/article/41218...
77	racetrack memory	2009	http://www2.technologyreview.com/article/41218...
78	software defined networking	2009	http://www2.technologyreview.com/article/41219...
79	traveling wave reactor	2009	http://www2.technologyreview.com/article/41218...
80	atomic magnetometers	2008	http://www2.technologyreview.com/article/40959...
81	cellulolytic Enzymes	2008	http://www2.technologyreview.com/article/40959...
82	connectomics	2008	http://www2.technologyreview.com/article/40959...
83	graphene transistors	2008	http://www2.technologyreview.com/article/40959...
84	modeling surprise	2008	http://www2.technologyreview.com/article/40959...
85	nanoradio	2008	http://www2.technologyreview.com/article/40959...
86	offline web applications	2008	http://www2.technologyreview.com/article/40959...
87	probabilisitic chips	2008	http://www2.technologyreview.com/article/40959...
88	reality mining	2008	http://www2.technologyreview.com/article/40959...
89	wireless power	2008	http://www2.technologyreview.com/article/40959...

90	a new focus for light	2007	http://www2.technologyreview.com/article/40747...
91	augmented reality	2007	http://www2.technologyreview.com/article/40747...
92	digital imaging reimagined	2007	http://www2.technologyreview.com/article/40747...
93	invisible revolution	2007	http://www2.technologyreview.com/article/40747...
94	nanocharging solar	2007	http://www2.technologyreview.com/article/40747...
95	nanohealing	2007	http://www2.technologyreview.com/article/40747...
96	neuron control	2007	http://www2.technologyreview.com/article/40747...
97	peering into video's future	2007	http://www2.technologyreview.com/article/40746...
98	personalized medical monitors	2007	http://www2.technologyreview.com/article/40747...
99	single-cell analysis	2007	http://www2.technologyreview.com/article/40747...
100	cognitive radio	2006	http://www2.technologyreview.com/article/40552...
101	comparative interactomics	2006	http://www2.technologyreview.com/article/40557...
102	diffusion tensor imaging	2006	http://www2.technologyreview.com/article/40552...
103	epigenetics	2006	http://www2.technologyreview.com/article/40552...
104	nanobiomechanics	2006	http://www2.technologyreview.com/article/40552...
105	nanomedicine	2006	http://www2.technologyreview.com/article/40552...
106	nuclear reprogramming	2006	http://www2.technologyreview.com/article/40552...
107	pervasive wireless	2006	http://www2.technologyreview.com/article/40552...
108	stretchable silicon	2006	http://www2.technologyreview.com/article/40552...
109	universal authentication	2006	http://www2.technologyreview.com/article/40552...
110	airborne networks	2005	http://www2.technologyreview.com/news/404001/1...
111	bacterial factories	2005	http://www2.technologyreview.com/news/404001/1...
112	biomechanics	2005	http://www2.technologyreview.com/news/404001/1...
113	cell-phone viruses	2005	http://www2.technologyreview.com/news/404001/1...
114	enviromatics	2005	http://www2.technologyreview.com/news/404001/1...
115	magnetic-resonance force microscopy	2005	http://www2.technologyreview.com/news/404001/1...
116	metabolomics	2005	http://www2.technologyreview.com/news/404001/1...
117	quantum wires	2005	http://www2.technologyreview.com/news/404001/1...
118	silicon photonics	2005	http://www2.technologyreview.com/news/404001/1...
119	universal memory	2005	http://www2.technologyreview.com/news/404001/1...
120	bayesian machine learning	2004	http://www2.technologyreview.com/featured-stor...
121	distributed storage	2004	http://www2.technologyreview.com/featured-stor...
122	microfluidic optical fibers	2004	http://www2.technologyreview.com/featured-stor...
123	nanowires	2004	http://www2.technologyreview.com/featured-stor...
124	personal genomics	2004	http://www2.technologyreview.com/featured-stor...

125	power grid control	2004	http://www2.technologyreview.com/featured-stor...
126	rNAi interference	2004	http://www2.technologyreview.com/featured-stor...
127	synthetic biology	2004	http://www2.technologyreview.com/featured-stor...
128	t-rays	2004	http://www2.technologyreview.com/featured-stor...
129	universal translation	2004	http://www2.technologyreview.com/featured-stor...
130	glycomics	2003	http://www2.technologyreview.com/featured-stor...
131	grid computing	2003	http://www2.technologyreview.com/featured-stor...
132	injectible tissue engineering	2003	http://www2.technologyreview.com/featured-stor...
133	mechatronics	2003	http://www2.technologyreview.com/featured-stor...
134	molecular imaging	2003	http://www2.technologyreview.com/featured-stor...
135	nanimprint lithogrpahy	2003	http://www2.technologyreview.com/featured-stor...
136	nanosolar cells	2003	http://www2.technologyreview.com/featured-stor...
137	quantum cryptogrpahy	2003	http://www2.technologyreview.com/featured-stor...
138	sooftware assurance	2003	http://www2.technologyreview.com/featured-stor...
139	wireless sensor networks	2003	http://www2.technologyreview.com/featured-stor...
140	biometrics	2001	http://www2.technologyreview.com/news/401765/b...
141	brain machine interface	2001	http://www2.technologyreview.com/news/400879/b...
142	data mining	2001	http://www2.technologyreview.com/news/401766/d...
143	digital rights management	2001	http://www2.technologyreview.com/news/401771/d...
144	flexible transistors	2001	http://www2.technologyreview.com/news/400877/f...
145	microfluidics	2001	http://www2.technologyreview.com/news/401770/m...
146	microphotonics	2001	http://www2.technologyreview.com/news/401768/m...
147	natural language processing	2001	http://www2.technologyreview.com/news/401767/n...
148	robot design	2001	http://www2.technologyreview.com/news/401769/r...
149	untangling code	2001	http://www2.technologyreview.com/news/400878/u...

Appendix B

Descriptions of Technologies and Companies

B.1 Game 1

Companies [46]:

- SolarCity: is a provider of solar energy services to homeowners, businesses, government and non-profit organizations, founded in July, 2006 and went public on the NASDAQ exchange in 2012 with a market cap of around $600 million.
- Siemens: is a multinational engineering and electronics company involved in the fields of industry, energy, transportation and healthcare. The company also provides financial products and services as well as insurance solutions. It operates in 190 countries and reported a global revenue of approximately 73.5 billion euros for the year of 2011. Siemens went public in March, 2001.

Starting technologies:

- SolarCity Gigafactory - Cluster 2 (Energy): SolarCity Gigafactory is a $750 million production production plant located in Buffalo NY. The solar facility will produce high efficiency solar panels amounting to a gigawatt of solar panels per year. When completed, it will be the largest solar manufacturing plant in North America and one of the biggest worldwide.
- Ultra-efficient solar power - Cluster 2 (Energy): Ultra-efficient solar power is enabled by a device that produces more than twice the solar power generated by today's panels by having an efficiency of at least 50 percent (compared to 20 percent today). The design allows the the sunlight to be split into six to eight component wavelengths each - similarly to a prism. Each wavelength component is then absorbed by a cell made with the semiconductor that can absorb it with the highest efficiency.

Target technology: solar fuel belongs to Cluster 2 (Energy)- benefit is 5.32: Solar fuel is a technology that uses sunlight to efficiently convert carbon dioxide into ethanol or diesel. This technology is based on the principle that bio fuels can be directly generated from carbon dioxide and water and the use of biomass such as corn or switch grass or algae can be eliminated since they are an intermediate step. Solar fuel achieves this by manipulating genes to create photosynthetic microorganisms [45].

Recommended Paths:

- For SolarCity ['SolarCity Gigafactory', 'supergrids', 'solar fuel']: cost is 1.5
- for Siemens ['ultra-efficient solar power 1','nanosolar cells','supergrids','solar fuel']: cost is 1.66

B.2 Game 2

Companies [46]:

- SolarCity was described in Game 1.
- Google is part of Alphabet, and is a multinational corporation that is specialized in internet-related services and products. The company offers an open-source mobile software platform and hardware products. Google X, is a semi-secret facility run by Google where scientists and entrepreneurs aim to improve technologies by a factor of 10, and to develop "science fiction-sounding solutions." The projects include; Project Glass, the Google driveless car, Project Loon, and Google Contact Lens.

Starting technologies:

- SolarCity Gigafactory is in Cluster 2 (Energy) and was described in Game 1.
- internet dna is in Cluster 1 (Living Matter): Internet dna is a system for trading genetic information between hospitals. It started with a focus on children with rare mutations in a single genes but can be applied to other areas of medicine. This system is called MatchMaker Exchange, and helps automate the comparison of DNA from patients around the world thereby transforming medicine through large-scale comparisons of genomes.

Target technology: solar fuel belongs Cluster 2 (Energy) with a benefit of 5.32 was described in Game1.
Recommended Paths :

- For SolarCity ['SolarCity Gigafactory', 'supergrids', 'solar fuel']: cost is 1.5
- For Google ['internet dna','glycomics','cellulolytic Enzymes','solar fuel']: cost is 2

B.3 Game 3

Companies:

- Magic Leap: Commercializes a proprietary virtual reality wearable technology. The company is creating a new user experience that is call Cinematic Reality™, where virtual objects are indistinguishable from reality.

- Google was described in Game 2.

Starting technology:

- magic leap is in Cluster 4 (Information). Magic leap is a virtual reality technology that uses a device to make virtual objects appear in real life. This technology is revolutionary for a number of domains especially in the film and gaming industries. The device uses a small projector that shines light onto a transparent lens. The reflected light has a pattern that blends with the surrounding light. This process tricks the visual cortex in a way that makes artificial objects indistinguishable from real objects [44].

- internet dna is in Cluster 1 (Living Matter) and was described in Game 2.

The target technology is solar fuel, it belongs to Cluster 2 (Energy) with a benefit of 5.32.
Recommended Paths:

- For Magic Leap ['magic leap','oculus rift','smart watches','personalized medical monitors','supergrids','solar fuel'] with a cost of 5.

- For Google ['internet dna','glycomics','cellulolytic Enzymes','solar fuel'],'alternatives_list': [['internet dna','glycomics','cellulolytic Enzymes','solar fuel'] with a cost of 2.

-

B.4 Game 4

Companies [46]:

- Apple is a multinational corporation that designs and manufactures electronics and software. It has a dominant market presence in the mobile communication and media devices, personal computers, it was founded by Steven P. Jobs, Steve Wozniak, and Ronald G. Wayne in April 1976. It went Public in December 1980. Its market capitalization is valued at $519.12B. Apple has recently started an effort to build a driveless car.

- Toyota is an automobile manufacturer operating 52 overseas manufacturing subsidiaries in 27 countries and regions. Toyota's vehicles are sold in more than and regions. Toyota manufactures a diverse line-up of vehicles sold in 160 countries. Toyota was founded in 1937.

Starting technologies:

- Brain machine interface is in Cluster 4 (Information) - Brain machine interface is a wireless brain-computer interface that enables patients to stream brain commands to the outside world via a wireless brain implant. Using a head-worn transmitter attached to a person's skull paralyzed patients can control objects such wheel chairs using electrical signals from the brain.

- solid state batteries is in Cluster 2 (Energy): Solid-state batteries are more compact, less flammable and more resilient than conventional lithium ion batteries. This is achieved by replacing the supporting materials (those that do not play a direct role in storing energy) inside the batteries with less cumbersome ones. For instance, while still using lithium ion technology, these batteries replace the liquid electrolyte with a thin layer of non flammable material. This technology is made possible by a simulation software that identifies the combinations of materials that can yield the right properties.

Target technology: car to car communication belongs to Cluster 3 (Information/Matter/Money) with a benefit of 3.30. Car-to-car communication enables cars to communicate with each other wirelessly within a few hundred meters. Each car can broadcast its position, speed, steering-wheel position, brake status, and other data to other vehicles. The availability of this data allows each car to build a picture of their environment and anticipate dangerous situations.

Recommended Paths:

- For Apple ['brain machine interface','deep learning','software assurance','synthetic biology','mechatronics','car to car communication']with a cost of 1.95

- For Toyota ['solid state batteries','nanoradio','mechatronics','car to car communication'] with a cost of 1.95

B.5 Game 5

Companies [46]:

- Apple was described in Game 4.

- Tesla Motors in an electric car company started by Elon Musk. Tesla has gone public in June 2010 and has a market cap of $2.47B. Tesla Motors' strategy is to first is to sell its own branded vehicle, second is to sell patented electric components to other automakers and to serve as a catalyst for the electric vehicle market.

Starting technologies:

- brain machine interface is in Cluster 4 (Information)was described in Game 4.

- Tesla Autopilot is in Cluster 2 (Energy): Tesla's autopilot is a software update delivered to the Tesla fleet of approximately 60,000 cars on the road at the time of the release. It enabled the embedded sensors to gather real-time data and use it for to manage speed, steering to change lanes, and park itself. While most of these features are present in other competitors' cars, the self-steering is the true advance toward driving autonomy.

Target technology: car to car communication belongs to Cluster 3 (Information/Matter/Money) with a benefit of 3.30 and was described in Game 4.

Recommended Paths:

- For Apple ['brain machine interface','deep learning','software assurance','synthetic biology','mechatronics','car to car communication'] with a cost of 1.95
- For Toyota ['Tesla Autopilot', 'car to car communication'] with a cost of 1.0

B.6 Game 6

Toyota: ['solid state batteries','liquid battery','universal memory','robot design','agile robots'] path length 2.58

Amazon: ['Conversational Interfaces','natural language processing','agile robots'] path length 2.0

Amazon also has three paths:

- Company to company with a shortest path length of 2: ['Amazon', 'Quake', 'Boston Dynamics']
- Technology to technology with a shortest path length of 2: ['Conversational Interfaces', 'natural language processing', 'agile robots']
- Company to technology with a shortest path length of 3: ['Amazon', 'F-Secure', 'Boston Dynamics', 'agile robots']

Bibliography

[1] Joseph L Bower and Clayton M Christensen. Disruptive technologies: catching the wave. Harvard Business Review Video, 1995.

[2] Chris Kimble and Isabelle Bourdon. The link among information technology, business models, and strategic breakthroughs: Examples from amazon, dell, and ebay. Global Business and Organizational Excellence, 33(1):58–68, 2013.

[3] Stephan Hess and Roland Yves Siegwart. R&d venture: proposition of a technology transfer concept for breakthrough technologies with r&d cooperation: A case study in the energy sector. The Journal of Technology Transfer, 38(2):153–179, 2013.

[4] Han TJ Smit and Lenos Trigeorgis. Strategic options and games in analysing dynamic technology investments. Long Range Planning, 40(1):84–114, 2007.

[5] Janghyeok Yoon, Hyunseok Park, Wonchul Seo, Jae-Min Lee, Byoung-youl Coh, and Jonghwa Kim. Technology opportunity discovery (tod) from existing technologies and products: A function-based tod framework. Technological Forecasting and Social Change, 100:153–167, 2015.

[6] Juneseuk Shin, Byoung-Youl Coh, and Changyong Lee. Robust future-oriented technology portfolios: Black–litterman approach. R&D Management, 43(5):409–419, 2013.

[7] Elicia Maine, VJ Thomas, and James Utterback. Radical innovation from the confluence of technologies: Innovation management strategies for the emerging nanobiotechnology industry. Journal of Engineering and Technology Management, 32:1–25, 2014.

[8] Bronwyn H Hall, Adam Jaffe, and Manuel Trajtenberg. Market value and patent citations. RAND Journal of economics, pages 16–38, 2005.

[9] Lionel Nesta and Pier Paolo Saviotti. Coherence of the knowledge base and the firm's innovative performance: evidence from the us pharmaceutical industry*. The Journal of Industrial Economics, 53(1):123–142, 2005.

[10] Rudi Bekkers and Arianna Martinelli. Knowledge positions in high-tech markets: Trajectories, standards, strategies and true innovators. Technological Forecasting and Social Change, 79(7):1192–1216, 2012.

[11] Eun Suk Suh, Michael R Furst, Kenneth J Mihalyov, and Olivier de Weck. Technology infusion for complex systems: A framework and case study. Systems Engineering, 13(2):186–203, 2010.

[12] Christopher Magee and Olivier de Weck. Complex system classification. International Council On Systems Engineering (INCOSE), 2004.

[13] Christopher L Magee and Oliver L de Weck. An attempt at complex system classification. In ESD Internal Symposium, 2002.

[14] Heebyung Koh and Christopher L Magee. A functional approach for studying technological progress: Application to information technology. Technological Forecasting and Social Change, 73(9):1061–1083, 2006.

[15] Sydney Do and Olivier de Weck. A grammar for encoding and synthesizing life support system architectures using the object process methodology. 44th International Conference on Environmental Systems, 2014.

[16] Hongyi Chen, Wayne Wakeland, and Jiang Yu. A two-stage technology foresight model with system dynamics simulation and its application in the chinese ict industry. Technological Forecasting and Social Change, 79(7):1254–1267, 2012.

[17] Nicolas Weidinger. 20 combinatorial forecasts map. http://www.iftf.org/our-work/people-technology/, 2014.

[18] Quid Team. Quid. https://quid.com/product/competitive-intelligence, 2016.

[19] Jean J Kong and Yue Kuen Kwok. Real options in strategic investment games between two asymmetric firms. European Journal of Operational Research, 181(2):967–985, 2007.

[20] Melanie Mitchell. Complex systems: Network thinking. SFI WORKING PAPER, (2006-10-036), 2006.

[21] ALBERT-LASZLO BARABASI. Network Science. Cambridge University Press, 2016.

[22] Albert-László Barabási and Jennifer Frangos. Linked: the new science of networks science of networks. Basic Books, 2014.

[23] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. science, 286(5439):509–512, 1999.

[24] Roger Guimera and Luis A Nunes Amaral. Functional cartography of complex metabolic networks. Nature, 433(7028):895–900, 2005.

[25] Tore Opsahl, Filip Agneessens, and John Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks, 32(3):245 – 251, 2010.

[26] KI Goh, ME Cusick, D Valle, B Childs, M Vidal, and AL Barabasi. Pnas. The human disease network. Department of Genetics, Harvard Medical School, 104(21):8685–8690, 2007.

[27] Dr. Cesar Ducruet and Dr. Jean-Paul Rodrigue. Graph theory: Measures and indices. https://people.hofstra.edu/geotrans/eng/methods/ch1m3en.html, 2016.

[28] David H. Reiley Jr. Avinash Dixit, Susan Skeath. Games of Strategy. BW. W. Norton and Company; 3rd edition, 2009.

[29] McLennan Andrew M. McKelvey Richard D. and Turocy Theodore L. Gambit: Software tools for game theory. http://www.gambit-project.org/gambit13/index.html, 2014.

[30] AlchemyAPI. Alchemyapi. http://www.alchemyapi.com/, 2016.

[31] Wikipedia. Natural language processing. https://en.wikipedia.org/wiki/Natural-language-processing, 2016.

[32] Linked Data. Linked data - connect distributed data across the web. http://linkeddata.org/, 2014.

[33] Erjia Yan and Ying Ding. Applying centrality measures to impact analysis: A coauthorship network analysis. Journal of the American Society for Information Science and Technology, 60(10):2107–2118, 2009.

[34] NetworkX Developers. Networkx. networkx. lanl. gov, 2010.

[35] Editor MIT Technology Review. Mit technology review. https://www.technologyreview.com/, 2016.

[36] Multiple. List of emerging technologies. https://en.wikipedia.org/, 2016.

[37] Multiple. Top 10 emerging technologies for 2014. https://www.weforum.org/agenda/2014/09/, 2014.

[38] Multiple. Introducing the technology pioneers 2015. http://widgets.weforum.org/techpioneers-2015b/, 2015.

[39] Hall Bronwyn, Adam Jaffe, Manuel Trajtenberg, et al. The nber patent-citations data file: Lessons, insights, and methodological tools. Patents, Citations, and Innovations: A Window on the Knowledge Economy, pages 403–460, 2001.

[40] Murat Bengisu and Ramzi Nekhili. Forecasting emerging technologies with the aid of science and technology databases. Technological Forecasting and Social Change, 73(7):835–844, 2006.

[41] TextBlob Developers. Textblob. https://textblob.readthedocs.io/en/latest/classifiers.html, 2016.

[42] Tom M. Mitchell. Chapter 3: Generative and discriminative classifiers: Naive bayes and logistic regression. Machine Learning, second edition, 2016.

[43] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment, 2008(10):P10008, 2008.

[44] Rachel Metz. Magic leap. MIT Technology Review, 2015.

[45] Kevin Bullis. Solar fuel. MIT Technology Review, 2010.

[46] Editor Tech Crunch. Tech crunch. Tech Crunch, 2016.

[47] Editor Thomson Reuters. The 2016 state of self-driving automotive innovation. Thomson Reuters, 2016.