Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών

Στρατηγικές 'Επιρροής-Εκμετάλλευσης' για Κοινωνικά Δίκτυα

Διπλωματική Εργασία

του

Συμινελάκη Παρασκευά

Επιβλέπων: Δημήτρης Φωτάκης
Λέκτορας Ε.Μ.Π.

Εργαστήριο Λογικής και Επιστήμης Υπολογισμών
Αθήνα, Μάρτιος 2012

National Technical University of Athens
School of Electrical and Computer Engineering
Division of Computer Science

Influence and Exploit Strategies for Social Networks

Undergraduate Thesis

by

Paris Siminelakis

Advisor: Dimitris Fotakis
Lecturer NTUA

Computation and Reasoning Laboratory
Athens, March 2012

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών

Στρατηγικές ‘Επιρροής-Εκμετάλλευσης’ για Κοινωνικά Δίκτυα

Διπλωματική Εργασία

του

Συμινελάκη Παρασκευά

Επιβλέπων: Δημήτρης Φωτάκης
Λέκτορας Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 2^{η} Μαρτίου 2011.

....	....	....
Ευστάθιος Ζάχος	Άρης Παγουρτζής	Δημήτρης Φωτάκης
Καθηγητής Ε.Μ.Π.	Επίκουρος Καθηγητής Ε.Μ.Π.	Λέκτορας Ε.Μ.Π.

Εργαστήριο Λογικής και Επιστήμης Υπολογισμών
Αθήνα, Μάρτιος 2012

(Υπογραφή)

....

Παρασκευάς Συμινελάκης

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Ευχαριστίες

Με το πέρας αυτής της διπλωματικής εργασίας κλείνει και ένας (ευχάριστος)κύκλος της ζωής μου. Ωστόσο κάποια πράγματα εξακολουθούν να υπάρχουν. Συγκεκριμένα, ένας ευρύτερος τρόπος σκέψης μαζί με κάποιες τεχνικές, οι ευχάριστες αναμνήσεις, οι ουσιαστικές ανθρώπινες σχέσεις αλλά και η ευγνωμοσύνη για τους ανθρώπους που με βοήθησαν και μου στάθηκαν.

Επιθυμώ , λοιπόν, να ευχαριστήσω πρώτα από όλους τον επιβλέπων καθηγητή μου Δημήτρη Φωτάκη για την αμέριστη συμπαράσταση και καθοδήγηση του. Όχι μόνο με μύησε στην νοοτροπία και πρακτική της θεωρητικής έρευνας αλλά ήταν και η αιτία, μέσα από την ενθουσιώδη διδασκαλία και αγάπη του για το αντικείμενο του, να ανακαλύψω τον δικό μου δρόμο και ερευνητικά ενδιαφέροντα. Οι συμβουλές του σε κρίσιμα στάδια της πορείας μου και η δικιά μου σύνεση να τις ακολουθήσω έχουν ουσιαστική συμβολή στην ακαδημαϊκή μου εξέλιξή. Θα ήθελα, επίσης, να ευχαριστήσω τον καθηγητή Στάθη Ζάχο, ο οποίος αποτελεί παράδειγμα ανθρώπου που μέσα από δύσκολες καταστάσεις εκτελεί το καθήκον του με συνέπεια, χιούμορ και ανθρωπιά. Τις ευχαριστίες μου έχουν και όλα τα μέλη του Εργαστηρίου Λογικής και Επιστήμης των Υπολογιστών(CoReLab), καθώς ο καθένας με τον δικό του τρόπο συνετέλεσε στην δημιουργία ενός ευχάριστου και φιλικού κλίματος συνεργασίας και συνύπαρξης.

Ιδιαίτερα ευχαριστώ τους στενούς μου φίλους και συμφοιτητές Τάσο Ουλή, Στέφανο Μπάρο, Θοδωρή Ντούσκα, Κώστα Παπασπύρου, Μυρτώ και Νικηφόρο Βλαχιά που μοιράστηκαν αυτά τα χρόνια του Πολυτεχνείου μαζί μου. Εύχομαι, ειλικρινά, να συνεχίσουμε να βλεπόμαστε μετά από χρόνια και να αναπολούμε τις στιγμές αυτές. Επίσης, δεν ξεχνάω και τους συμφοιτητές-φίλους Γιάννη Τσιούρη, Χρήστο Σταυρακάκη τους οποίους πέρα από στιγμές γέλιου χρωστάω και μία εργασία στις βάσεις ;).

Τέλος, θέλω να ευχαριστήσω την οικογένεια μου για την υποστήριξη, παρότρυνση και πίστη τους σε εμένα. Ιδιαίτερα, θέλω να αναφέρω την γιαγιά μου Ουρανία Ηλιοπούλου, η οποία είναι παράδειγμα αυτοθυσίας και προσφοράς και συνέβαλε σημαντικά στο να αγαπήσω την Γνώση.

Πάρης Συμινελάκης

Περίληψη

Η Κοινωνική Δικτύωση μέσω Ιντερνέτ έχει αποκτήσει κεντρική θέση για την Διαφήμιση και την Εξόρυξη Δεδομένων για εμπορικούς σκοπούς. Εταιρείες Κοινωνικής Δικτύωσης(π.χ. Facebook, Orkut, Google+) διατηρούν λεπτομερή δεδομένα για εκατομμύρια χρήστες, τα οποία τα διαθέτουν σε εταιρείες για να αυξήσουν την διεισδυτικότητα των προϊόντων τους στην αγορά. Τα έσοδα απο διαφημίσεις των δικτύων αυτών αποτελούν την βάση ενός βιώσιμου επιχειρηματικού μοντέλου και χρησιμοποιούνται για να υποστηρίξουν τις τεχνικές υποδομές και να διασφαλίσουν την ποιότητα των υπηρεσιών τους. Ωστόσο, υπάρχει μεγάλη διαφοροποίηση μεταξύ των πραγματικών εσόδων και της εκτιμούμενης αξίας των εταιρειών αυτών. Είναι κοινός τόπος ότι πολλές από τις δυνατότητες των Εταιρειών Κοινωνικών Δικτύων παραμένουν αναξοιοποίητες και η πεποίθηση αυτή έχει συνοδευτεί από εντατικές ερευνητικές προσπάθειες για την Εμπορευματοποίηση των Δεδομένων από Κοινωνικά Δίκτυα.

Στόχος της διπλωματικής εργασίας είναι η κατανόηση και επέκταση των τεχνικών για την αξιοποίηση της γνώσης του ιστού των κοινωνικών σχέσεων και των μεταξύ τους αλληλεπιδράσεων. Πραγματοποιείται ανασκόπηση της βιβλιογραφίας και επικεντρωνόμαστε σε δύο σημαντικά και στενά σχετιζόμενα προβλήματα, στο πρόβλημα Μεγιστοποίησης της Επιρροής[Kempe, Kleinberg, Tardos'03] και στο πρόβλημα της Μεγιστοποίησης των Εσόδων[Hartline, Mirrokni, Sundararajan, '08]. Το πρόβλημα Μεγιστοποίησης της Επιρροής πραγματεύεται περιπτώσεις όπου άνθρωποι καλούνται να πάρουν μια δυαδική απόφαση(αγοράσουν ένα προϊόν, ψηφίσουν ένα υποψήφιο, υιοθετήσουν μια καινούργια τεχνολογία) και ζητειται το βέλτιστο αρχικό σύνολο ανθρώπων δεδομένου μεγέθους που μέσω της επιρροής τους θα οδηγήσουν στην μέγιστη δυνατή διάδοση. Το πρόβλημα της Μεγιστοποίησης των Εσόδων αφορά την σχεδίαση στρατηγικών πώλησης προϊόντων, των οποίων η αξία για κάθε αγοραστή αυξάνει ανάλογα με το ποιοί γνωστοί του ήδη το κατέχουν, εκμεταλλευόμενοι την γνώση του κοινωνικού ιστού. Εστιάζουμε την προσοχή μας σε μια κλάση στρατηγικών 'Επιρροής και Εκμετάλλευσης¨(ΕΕ), όπου ένα αρχικό σύνολο ανθρώπων έχουν ευνοϊκή μεταχείρηση(δωρεάν δείγματα, χρηματικά ανταλλάγματα) ώστε να κερδίσουμε την επιρροή τους στο δίκτυο και οι υπόλοιποι αντιμετωπίζονται με τρόπο ώστε να πετύχουμε τον στόχο μας(υψηλότερα έσοδα, μεγαλύτερη αποδοχή).

Η τεχνικής συνεισφορά της διπλωματικής εργασίας αφορά το Πρόβλημα Μεγιστοποίησης Εσόδων υπό το Ομοιόμορφο Αθροιστικό Μοντέλο[Hartline et al.'08]. Αρχικά αποδεικνύουμε ότι το πρόβλημα είναι NP-Δύσκολο ακόμη και όταν το δίκτυο δεν είναι κατευθυνόμενο, χρησιμοποιώντας μια αναγωγή από το πρόβλημα Monotone One-in-Three SAT. Στην συνέχεια πραγαματοποιούμε μια συστηματική διερεύνηση των αλγοριθμικών ιδιοτήτων των στρατηγικών 'Επιρροής-Εκμετάλλευσης¨. Αποδεικνύουμε ότι το πρόβλημα σχεδιασμού της Βέλτιστης στρατηγικής 'ΕΕ' είναι NP-Δύσκολο και παρέχουμε ένα κάτω φράγμα για τον λόγο των εσόδων απο μια τέτοια στρατηγική και των μέγιστων δυνατών εσόδων. Επιπρόσθετα, επεκτείνουμε και βελτιώνουμε την απλή στρατηγική 'ΕΕ' των Hartline et al., βελτιώνοντας κατα λίγο τον λόγο προσέγγισης του προβλήματος. Η κύρια συνεισφορά έγκειται στην σχεδίαση στρατηγικών 'ΕΕ' βασιζόμενοι σε Ημιορισμένες Μεθόδους Χαλάρωσης Ακέραιων Προγραμμάτων και η ακόλουθη σημαντική βελτίωση που επιτυγχάνεται στον λόγο προσέγγισης της βέλτιστης λύσης. Τέλος, προτείνουμε μια οικογένεια στρατηγικών Τοπικής Αναζήτησης για την βελτίωση μιας οποιαδήποτε λύσης καθώς και Ευριστικές Μεθόδους βασιζμένες σε Ιδιοδιανύσματα για την συσχέτιση της θέσης ενός ατόμου στο δίκτυο και την τιμή που θα του προσφέρουμε.

Λέξεις Κλειδιά

Στρατηγικές 'Επιρροής-Εκμετάλλευσης¨, Μεγιστοποίηση Επιρροής, Μεγιστοποίηση Εσόδων, Προσεγγιστικοί Αλγόριθμοι, Εμπορευματοπίηση Κοινωνικών Δικτύων, Θετικά

Abstract

The importance of online social networks in advertising and market research is by now indubitable. Social networks provide detailed and broad information for millions of users and companies have been using this information to increase market penetration of their products. Social network companies use the revenue exerted by advertisements to sustain the costs involved in maintaining their servers and quality of service, as well as to provide the basis of a sustainable business model. However, there is a large discrepancy between the perceived value of Social Networks and the actual revenue they generate.The widespread belief is that much of the potential of Social Networks remains unexploited. This premise has spurred a large amount of research in the direction of mon- etizing Social Networks.

In this thesis, we are concerned with utilizing the information about the structure and strength of social ties in order to achieve certain objectives. We review previous approaches and focus on two important and closely related problems, that of Influence Maximization[Kempe, Kleinberg, Tardos'03] and Revenue Maximization[Hartline, Mirrokni, Sundararajan, '08]. The Influence Maximization Problem considers situations where a binary decision is made about adopting or not an innovation(product,technology,behaviour) and seeks for the best seed of initial adopters that achieve overall maximum spread by interacting with their social contacts. On the other hand, the Revenue Maximization Problem aims at exploiting positive network effects between buyers about the value of a product to devise a marketing strategy that maximizes the revenue. We focus an a class of strategies called Influence and Exploit, where a set of individuals is treated preferentially(free product, monetary incentives) in order to "seed" the network(Influence) and then the remaining individuals are exploited(full price, no incentives) to achieve the objective(higher revenue, wider adoption).

The technical contribution of this thesis concerns the Revenue Maximization Problem under the Uniform Additive Model[Hartline et al.'08]. We initially prove that the problem remains NP-Hard even for the undirected case via a reduction from Monotone One-in-Three SAT. Then, we embark a systematic study of the algorithmic properties of Influence and Exploit strategies. We prove that finding the Optimal Influence and Exploit strategy is NP-Hard and provide lower bounds on the ratio between the revenue extracted from an optimal IE strategy and the optimal revenue in general. Furthermore, we slightly extend and optimize the simple IE strategies proposed by Hartline et. al obtaining a first improvement of the approximation ratio of the problem. Our main technical contribution lies in developing powerful Semidefinite Programming Relaxations for designing IE strategies and the corresponding significant improvement on the approximation ratio for the problem. Finally, we propose a class of Local Search strategies to improve on an given solution and introduce intelligent heuristics based on Eigenvector Centrality correlating explicitly network position and the price to be offered to each buyer.

Keywords

Influence and Exploit, Influence Maximization, Revenue Maximization, Approximation Algorithms, Social Network Monetization, Positive Network Externalities

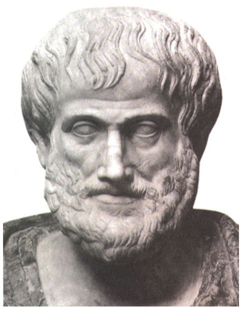
Contents

Chapter 1

Introduction

This thesis is part of the greater endeavour of Computer Science to understand social phenomena and processes. We first present arguments on why are social phenomena important and what are the benefits from strengthening our understanding of them. Then we describe the unique position that Computer science finds itself today in this endeavour.

The Greek Philosopher Aristotle [7] believed that the true purpose of man is the attainment of one's eudaimonia. He defined eudaimonia as the cultivation of virtue and the realization of one's true potential. He argued that "man is a social being" and an essential condition towards that goal is friendship. Actually, we argue not only that man is by nature a social being but something stronger; it is exactly the social nature of man that is responsible, at large, for man as we perceive him today. The organization of humans in groups and societies has facilitated, enabled and inspired the greater achievements of our civilization: *division of labour, language, reason, science*; in one word *progress*.



Already from the early years of human presence on earth, people have formed groups. Initially, the basic unit would be a small "pack", where there would be a coarse grain division of people in working groups depending on gender and age. Gradually, starting with the invention of agriculture[43], people formed larger communities and a form of hierarchic authority began to emerge. Nowadays, the social and political organization has a self-similar, almost fractal, hierarchical organization. Viewing the social structure as networks, we began with small, almost disjoint, tightly knit networks. These networks merged under a common authority and formed larger networks with power flowing in star network topology. This process was repeated at many layers and has given rise to the complex socio-political structure that we see today.

Unsurprisingly, the evolution of human civilization is closely related with the evolution of social organization. The increasing complexity of social organization has brought three main classes of evolutionary advantages concerning: *performance, diffusion, robustness.* Communities increased their survivability and standards of living by assigning roles to its members, according to their own special characteristics and talents(strength, dexterity, intelligence, upbringing capabilities). This way individual members or families did not need to be autonomous and by depending on others increased their performance at fulfilling their individual duties. Another important factor for the survival and progress of a community is in what scale and how fast individual knowledge about the world becomes available to other members. This knowledge could be about potential hazards (fire, war, etc.), opportunities or even technology

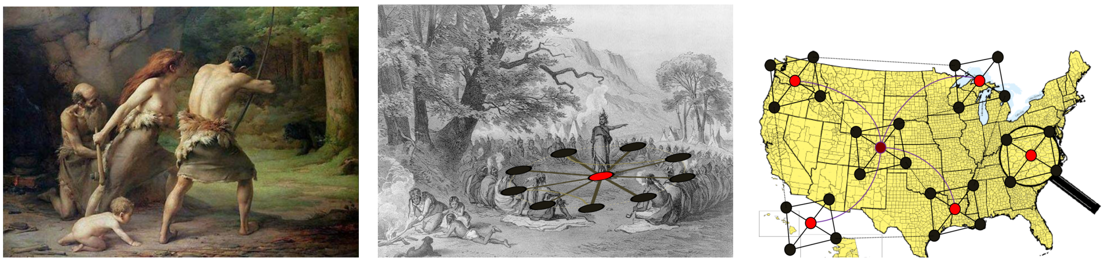


Figure 1.1: The historical evolution of social networks.(a)Prehistoric Family(*Prehistoric Man Hunting Bears*, Emmanuel Bener, Musee d'Unterlinden, Colmar, France).(b)Indian tribe council.(c)State and County organization of US.

and ideas. Naturally, communities where this diffusion is more efficient would be more fit to respond to changes of the environment. Finally, one critical characteristic for the survival of both a living organism and a society is the robustness under random or even targeted failures. Imagine that for some reason(e.g. battle, illness) the top authority of a community passed away; if there weren't people fit to replace him or to respond appropriately until the new leader is selected, the community would be vulnerable either to hostile communities or to internal strife. The same reasoning applies for other key people in a community, such as a doctor, a maid or a strong warrior. In a way natural selection has been optimizing the structure of social organization weighing appropriately all three factors.

One great thing about the human species is that it can circumvent the slow painstaking process of natural selection via scientific research. Traditionally, sciences such as sociology, anthropology and management science have acknowledged the important role that social phenomena play and have proceeded with comprehensive studies aiming at unravelling the impact of social interaction in a variety of settings[101, 115, 63, 124]. Therefore, these sciences have accelerated developments concerning social organization as their research has been taken into account in the decision-making process of governments, organizations and companies. However recent technological advances and globalization have brought about a need for a new approach.

Today social interactions have been revolutionized by the explosion in Transportation, Telecommunications and Information Technology. The phone, cell phone and above all the Internet have changed forever the way people interact. Social interaction has crossed the spatial and national boundaries and nowadays people communicate with each other across countries, continents and cultures. Therefore, old scientific methods of studying social phenomena using questionnaires and searching public files are almost obsolete. Nevertheless, there is a new weapon in our arsenal; the Internet.

The internet is a place where ideas, data and products are being exchanged. Besides its profound value for the scientific community, the Internet has also great impact on the economy. The rise of corporate giants as Google and Facebook are indicative of its potential. Moreover, the value of e-commerce and online retailing in U.S. alone is estimated to be up to 200 billion dollars (Forrester Research 2011). Perhaps, what is more interesting about the Internet is the fact that we have for the first time a clear account of the transactions taking place, as most information is recorded and theoretically could become available for study and analysis.

An important recent artifact of the Internet is the appearance of Social Networks. Social Networks are web-sites or applications where users subscribe and interact with other users. They can share photos, music, video, opinions, easily chat or even have multi-person conver-

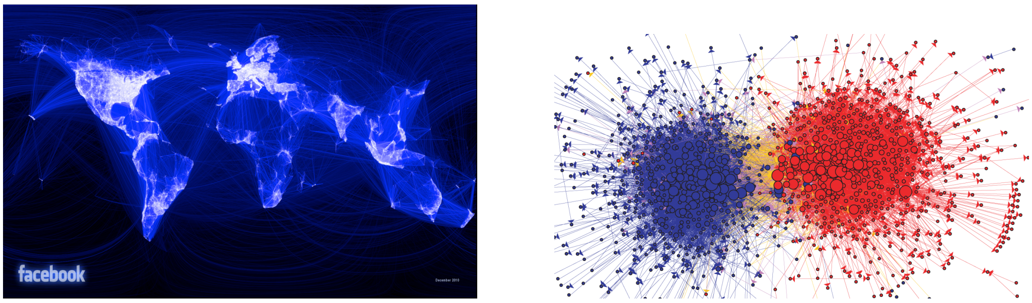


Figure 1.2: The increasing complexity of modern social networks.(a) World Map and Friendship network of Facebook. (*Visualizing Facebook Friends*, Paul Butler).(b) *The Political Blogosphere and the 2004 U.S. Election*, Lada Adamic, Natalie Glance

sations. Social Networks have great value for the economy and society as, primarily, they cover the basic need for communication and sharing. They facilitate the diffusion of ideas, news, products etc. and, crucially for science, they record the social interactions taking place between members. For the first time we can actually peer into the actual network of friendship simultaneously for thousands or million of people, fact that sociologists could not even dream of. The sheer size and amount of data that have become combined with digitized information have provided the impetus for Computer Science to get involved with the study of social networks.

1.1 Diffusion of Innovations

A key question in Sociology and Economics is why some ideas, norms or products spread and become widely accepted, whereas others die out. Is it only a fact of their intrinsic value and appeal or are there other latent phenomena involved? The first systematic study towards this direction was carried out by Everret Rogers [115] who unified previous approaches into a concrete theory, coined with the term *Diffusion of Innovations*.

Diffusion of Innovations is a theory seeking to explain how, why and at which rate do ideas, technology and products spread through societies. Rogers identifies four main elements that influence the spread: the *innovation* itself, the *communication* channel, i.e. the medium through which messages and information about the innovation spread, *time*, since the decisions made by individuals have a strong temporal dependence, and, finally, the *social system* which consists of all the interested parties(individuals) involved in the decision process.

The mechanism of diffusion is thought to occur through a five-step decision process, which is mediated by a series of communication channels over a period of time among members of the social system. The first stage is *knowledge*, where the individual is initially exposed to the innovation but has incomplete information about it. Then follows *persuasion*, a stage during which the individual is interested in the innovation and actively seeks more information. The next stage is *decision*. At this stage the individual weighs the relative benefits and costs of adopting the innovation and makes a decision. If the individual decides to adopt the innovation he proceeds to a stage called *implementation*, where he employs the innovation to a varying degree and may seek further information about it. The last stage is *confirmation* upon which the adoption of innovation is finalized after the first trial period has produced positive results and the innovation is used to its fullest potential.

Rogers continues his investigation and defines several intrinsic characteristics of innovations. The *relative advantage* of an innovation with respect to the previous generation, the *complexity* and *trialability* of the innovation, that is how steep is the "learning curve" associated with a given innovation, the level of *compatibility* with the individuals life, i.e. how easy is to integrate the innovation to one's life and the *observability*, which is the extent that an innovation is visible to others and is implicitly communicated. These all are factors that depending on the situation can have an important impact on the adoption or not of an innovation.

It has been empirically shown that the adoption of an innovation follows an S-curve when plotted over a length of time. In order to provide further insight into the diffusion process, Rogers also makes a classification of individuals within a social system into categories based on the degree of *innovativeness*. As an attempt to construct a "derivative" of the adoption curve, he considers that individuals are normally distributed with respect to the time they decide to adopt the innovation. He suggests five categories of adopters: innovators, early adopters, early majority, late majority and laggards. *Innovators* are the first individuals to adopt an innovation. They are usually young, risk-taking individuals of high social status and strong financial grounding with closest contact to scientific resources. *Early adopters* are the second fastest people to adopt an innovation. They usually have the highest degree of opinion leadership and are of the highest social class and command more financial resources. The *early majority* adopt an innovation after a varying length of time and have above average social status and contact with early adopters, which act as opinion leaders. The *late majority* will adopt the innovation after the average member of the society. They approach the innovation with a high degree of scepticism. They usually have bellow average social status and very little financial lucidity. The last people to adopt an innovation are the *laggards*. These people are change-averse and tend to be advanced in age. They are usually focused on traditions and are likely to have the lowest social status. So this categorization attempts to explain the logistic growth of the adoption process as a consequence of the wide distribution of the innovativeness characteristic within a population.

Despite the aggregate nature of the Diffusion of Innovations theory, Rogers acknowledges the existence of highly influential individuals in a social system and the impact that they can have on the final outcome of the adoption process. Relying on ideas of Katz and the notion of Centrality he develops his ideas about the role of Opinion Leadership. Moreover, he also raises the issue of the impact of the social structure in the diffusion process observing that homophily can hinder or amplify the spread of an innovation. *Homophily* is the tendency of people to socialize with other people that are similar in attributes such as age, nationality, occupation, social status etc. Homophilous individuals communicate more effectively and therefore their contact will lead to greater knowledge gain. However, homophily can be also an obstacle to the communication process since homophilous individuals tend to have the share information and new ideas are hard to be introduced. Therefore, an ideal situation would be for two individuals to be homophilous in every way, except the knowledge about the innovation. These early concerns already start to elucidate the importance of network topology in the diffusion of innovations through networks.

Another attempt to mathematically model the diffusion of innovations on an aggregate level was made by Bass[11]. He assumes that the population is entirely homogeneous, i.e. everyone is equally likely to interact-communicate with anyone else. In making their decisions, individuals are influenced by two sources. The first source is mass marketing such us advertising and the second is "word or mouth" effects between individuals who have already adopted the innovation. The Bass model traces the probability $R(t)$ that a random individual has adopted the innovation up to time t and because of the homogeneity of the population this corresponds to the ratio of the population that have adopted the innovation. Now for each t, a random individual has probability $1 - R(t)$ of not having adopted the innovation. This individual is convinced in the next time frame $t + dt$ from mass marketing with probability p. If that fails he has a probability $(1 - p)q$ to adopt the innovation if he meets another individual who already adopted the innovation which happens with probability $R(t)$. Putting all the pieces together we get that the differential equation describing the ratio of buyers that have adopted the product prior to time t is:

$$\frac{dR(t)}{dt} = (1 - R(t))[p + (1 - p) \cdot q \cdot R(t)]$$

Mathematically this is a Riccati equation with constant coefficients and can be solved for various parameters p, q. The model is widely used in product and technology forecasting and has been generalized to include other aspects as well[127]. The main assumption this model makes is that the probability of purchase is linearly related to the number of previous buyers, which in turn implies exponential growth of initial purchases and then exponential decay; behaviour which is typical of sigmoid curves.

1.2 Threshold Phenomena

In trying to explain collective behaviour, social sciences operated on the premise that when we observe a collective outcome we can infer that these individuals ended up sharing the same belief about the situation, even if they did not in the beginning. Granovveter [62] based on Schelling's Model for residential segregation [123] proposed models of collective behaviour that showed with the brightest colours how individual variations of norms can have an adverse effect on the final outcome.

His model treats situations were binary decisions are made(diffusion of innovations, riots, voting etc.) and the cost/benefit from either decision depends on how many individuals have followed each action. Individuals are assumed to be rational and make the decision that serves their interests best. In the simplest case of the model, each individual i chooses a threshold t_i and becomes becomes "active" only if t_i more individuals have already decided to be active as well. Consider for instance a peacefull protest that can potentially escalate into a riot. Assume there are 100 individuals and each individual chooses to participate in the riot if a certain number of people already participate. Assume that their thresholds are $t_1 = 0, \ldots, t_{100} = 99$. In this case all individuals will eventually be engaged in the riot. Whereas a very similar distribution of thresholds and thus beliefs where only t_2 would be 2 instead of 1, would result in only one single "demented" individual's violent acting. This example illustrates that it is fallacious to infer individual norms from collective behaviour and that aggregating individual beliefs can result in varying outcomes depending both on the distribution of beliefs as well as to the way they are communicated between individuals.

Already it is eminent that explaining and predicting collective behaviour is a complex problem. We started from aggregate models and theories which are exemplified by the Diffusion of Innovations theory of Rogers, and continued to models where individual behaviour is taken into account through uniform interactions. However, when considering the threshold models one can immediately sense that there is another latent factor that impacts greatly on the final outcome; network structure or the configuration of "social visibility". Morris[104] studied a setting where all individuals have the same threshold but they can only perceive the actions and beliefs of specific other individuals. He considers a local interaction game where

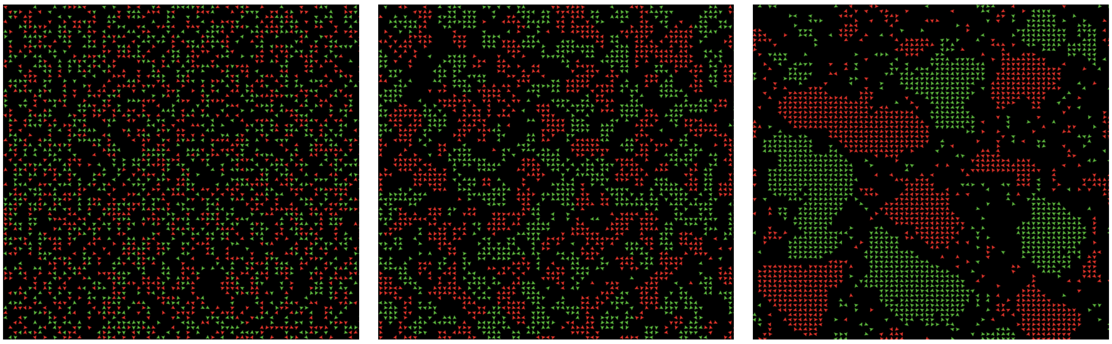


Figure 1.3: T.Schelling's model of Residential Seggregation. The residents deside to move when the number of neighbours of the same color drops below a threshold. This model shows that residential segreggation needs nots be a result of racism rather than a slight preference to be around with people of the same race. The three figures are the steady state of the model for different values of the threshold(50%,60%,72%)

each individual selects to follow one of two actions and receives a pay-off that depends on the fraction of his neighbours that follow each action. Particularly, every individual is modelled by a node and individuals that influence each other are connected by edges, i.e. they are "neighbours" in the social graph. The local interaction game is parametrized by a threshold $0 < q < 1$, according to which an individual switches to a behaviour when at least a q-fraction of his neighbours follow that action. Morris studied graphs with countably infinite nodes and seeks to find the *contagion threshold* of the graph, that is the smallest q such that a behaviour that is followed by some finite population can spread to the entire population. He provided contagion thresholds for various infinite lattices and studied further properties of such local interaction games. One important aspect that stemmed from his research is the impact of tightly knit communities, which can impede the spreading process.

1.3 Externalities

An important area of study in Economics is the issue of *externalities*. Externalities capture the idea that there are involuntary third party cost or benefits involved from a voluntary transaction or action between two parties. Classical examples are pollution, waste management or network externalities where the valuation of a product or technology depends on how widely used it is[74, 75]. There are three classical solutions to the problem of externalities aiming to improve decision making[128]:

- Coase Negotiation, where agents negotiate their way to an efficient outcome since the exploitation oft externalities provide sufficient economic incentive.
- Setting up a competitive market for the externality, for instance the right to pollute, where it is assumed that the market dynamics lead to an efficient outcome.
- Imposing a Pigovian Tax, where the externalities are imposed on the agents by a regulator through taxations(negative externalities) or economic subsidies(positive externalities).

The past years there is large amount of research in Computer Science unravelling the impact of externalities in a variety of settings. We briefly provide some indicative examples.

Combinatorial Auctions One of the main questions lying in the intersection of computer science and economics, is how to allocate a set of resources in order to maximize an objective(social welfare, revenue, etc.). Selling single items(resources) is done relatively easily, but how about selling multiple items where there substitutabilities and complementarities between items? This issue is handled in the study of Combinatorial Auctions, where there are complex(combinatorial) dependencies between items(for reviews see[37, 109]). The situation can become even more complex if we consider that bidders not only care about which combination of items they are allocated but also about which items other bidders get, namely the issue of externalities. Krysta et al.[86] first considered externalities for Combinatorial Auctions. They developed a bidding language, to succintly represent bidders valuations and study algorithmic propoerties of the winner determination problem and the complexity of characterizing bidders valuations. Conitzer, Sandholm[33, 34] and Lu[100] study settings with externalities, where an agent controls one or more variables and how these variables are set affects not only the agent herself, but also potentially the other agents. Furthermore, there are many studies considering the impact of negative externalities[23, 15, 35, 40, **?**].

Congestion Games Blumrosen and Dobzinski[18] consider Congestion games, where there is a set of resources and players select a subset of them acting selfishly, under various models of externalities(positive,negative), that is different cost models for sharing a resources. They provide algorithmic relations with Combinatorial Auctions allowing to translate techniques and intuition from that field to centralized Congestion Games. The authors design constant approximation algorithms for Welfare Maximization and provide hardness results under both positive and negative externalities. Furthermore, they construct an $O(\sqrt{n})$-approximation mechanism and show how to compute "Order-Preserving" equilibria in polynomial time. They also show that for non-anonymous(player specific externalities) congestion games only trivial approximations can be guaranteed. Fotakis et al.[47] study externalities in Congestion Games from the perspective of incomplete information, that is individual's valuation only depend on the strategy of other neighbouring buyers in the social graph. They show that such games admit a potential function and therefore have a Pure Nash Equilibrium(PNE). They provide results that explicitly quantify the Price of Anarchy(PoA) and Price of Stability(PoS) in terms of the independence number of the graph. Finally, they study the time needed for ϵ-Nash dynamics to reach an approximate equilibrium.

Sponsored Search Auctions The main source of revenue for web services companies(search engines, social networks) is online advertising. In the past, advertising was conducted through static banners on websites renting their space for a limited amount of time. The last decade a new method of advertising has prevailed, that of Sponsored Search Auctions. Web search engines monetize their service by auctioning off advertising space next to their search results. That is, there are a limited amount of slots available for advertisements next to each search result and the advertisers make bids to occupy them. This subject has received extensive attention in the computer science community(for reviews see [109, 133]), however the issue of externalities has been considered only recently. Externalities were first considered by Ghosh and Mahdian[**?**] after there were experimental evidence for the hypothesis that the click-through rate of ads depend on surrounding ads by Joachims et al.[72]. Various models of externalities and the corresponding problem of winner determination have been studied since[77, 1, 48] as well as properties of Nash Equilibria for GSP mechanisms used in practice[53, 41, 57, 48]. Furthermore, the value of learning and price of truthfulness in such mechanisms has also been investigated[42, 97].

1.4 Summary and Organization

In this thesis we focus on social interaction rather than economic or market dynamics. The central question is now that we have explicit knowledge for the social network: *How can we utilize the knowledge of the social graph?*. We are interesting in either specifying the impact of the social structure on social processes or how to devise mechanisms that facilitate desired events and outcomes.

We review previous approaches and focus on two important and closely related problems, that of Influence Maximization[Kempe, Kleinberg, Tardos'03] and Revenue Maximization[Hartline, Mirrokni, Sundararajan, '08]. The Influence Maximization Problem considers situations where a binary decision is made about adopting or not an innovation(product, technology, behaviour) and seeks for the best seed of initial adopters that achieve overall maximum spread by interacting with their social contacts. On the other hand, the Revenue Maximization Problem aims at exploiting positive network effects between buyers about the value of a product to devise a marketing strategy that maximizes the revenue. We focus an a class of strategies called Influence and Exploit, where a set of individuals is treated preferentially(free product, monetary incentives) in order to "seed" the network(Influence) and then the remaining individuals are exploited(full price, no incentives) to achieve the objective(higher revenue, wider adoption).

In what follows we summarize previous work done in the above problems and then we present our own original results[134]. Chapter 2 concerns the problem of Influence Maximization, where we briefly present the key results. In chapter 3, we summarize previous approaches to the problem of Revenue Maximization and emphasize the work of Hartline et.al[66] upon which we have based our research. In Chapter 4, we provide insights and hardness results about the model considered by Hartline et.al and then proceed in Chapter 5 where we design approximation strategies improving previous work. We accomplish that by using and extending two approaches; randomly partitioning vertices into pricing classes, inspired from Hartline et.al, and randomized rounding of a semidefinite program, as pioneered by Goemans and Williamson. In Chapters 6 , we provide Local Search strategies, to improve a solution, and intelligent Heuristics correlating the right price to be offered to a buyer with his network position. Finally, in Chapter 7 we discuss the validity of the model as well as other issues.

Chapter 2

Influence Maximization

The study of social processes by which ideas and innovations diffuse through social networks has been ongoing for half a century and as a result a fair understanding of such processes has been achieved. Modern models about social influence have been augmented with various features allowing for arbitrary network structure, non-uniform interactions, probabilistic events and other aspects. Therefore, scientists have now turned to the next frontier which consists of two complementary directions: obtaining accurate estimates of the parameters involved(graph, weights, probabilities) and utilizing the knowledge available to guide those processes in order to meet certain objectives(e.g. wider diffusion, greater profits). In this chapter we will focus solely on the latter direction.

Traditionally, advertising would be conducted through a channel and individuals, with access to that channel, would be influenced at the time frames that the message is being transmitted . However, recent studies have shown that traditional channels are loosing their reach[90] and that custom-tailored messages are much more effective than traditional mass marketing. Moreover, many times personal recommendations from family, friends and co-workers have greater impact than any third party messages. Hence, the idea is instead of selecting or creating a channel upon which to transmit a message, to use the very social network of personal acquaintances as the medium. This approach presents the advantage of reduced costs, higher efficacy and almost universal reach. It is known with the popular term of Viral Marketing, since the message spreads like a virus between individuals that come into contact. The recent success of the Hotmail email service, which on a tiny budget of 50.000 reached 12 million users in 18 months only due to a promotional url at the end of the message, is indicative of the potential of Viral Marketing methods. Nevertheless, many companies have tried since to repeat the success and invest heavily on obtaining a customer basis only to see their product never actually to start off. Why is it that?

The success of viral marketing strategies depends heavily on our ability to appropriately "seed" the network with an initial number of followers which will convey our message. Domingos and Richardson[44] modelled the underlying interactions between buyers in the stochastic framework of Random Markov Fields and posed the fundamental algorithmic question: If we knew the network of personal relationships how can we select an initial number of individuals to influence so that after the cascading process terminates the number of affected individuals is maximal? They provide heuristics that select individuals with a large effect on the network, develop techniques to extract the necessary influence data and conduct some experiments.

Kempe, Kleinberg and Tardos in their seminal paper *Maximizing the Spread of Influence through a Social Network*[76] formalized the question posed by Domingos and Richardson in

Figure 2.1: Some examples of threshold phenomena, where there is a binary decision to be made and actors influence each other.(a) Elections.(b)Rioting.

algorithmic terms as a discrete optimization problem, under the name of *Infleunce Maximization*. They consider a simple context where all individuals are initially *inactive*(not willing to buy our product, vote for our candidate or adopt our innovation). Then we select a limited number k of individuals to *influence*(e.g. monetary incentives, individualized marketing, free samples) so that they become *active*. They assume that individuals influence each other and that active individuals may "infect" other individuals to become active as well, according to a stochastic model of interaction.

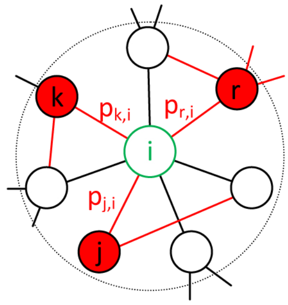
The problem, therefore, is given a social network, i.e. a set of nodes(individuals) and the edges(interactions) between them, to select the optimal "seed" of individuals to influence so that after the activation process terminates the number of active nodes is maximal for a seed of size k. KKT studied the most widely used influence models from sociology and interacting particle systems and showed the problem to be NP-Hard even for the simplest model. They then prove some crucial properties about these models and based on them utilize a result of Nemhauser et.al[107] to provide a greedy 0.63 approximation algorithm for a general class of models. They also conducted experiments showing that their algorithm outperforms heuristics based on centrality eigenvector methods from social network analysis.

2.1 Influence Models

We first describe the two basic influence models, namely the Independent Cascade Model and the Linear Threshold Model. We further provide some properties of these models and use them to derive results for the problem of Influence Maximization.

2.1.1 Independent Cascade Model(IC)

The ICM was introduced by Goldenberg et.al[56] to model the dynamics of viral marketing and is inspired from the field of interacting particle systems[99, 16]. In this model, we start with an initial set A_0 of active individuals, each active individual has a single chance to activate each non active neighbour of his. However, the process of activation is deemed stochastic and succeeds(fails) with constant probability q(respectively $1-q$) independently for each attempt. Therefore, from an initial population of active individuals the activation process spreads in a cascading manner as newly activated individuals may activate new nodes that either previous attempts failed to activate or were not before accessible.



There is an elegant interpretation of the ICM, in terms of the reachability of nodes via paths from the initial active set A_0. We can picture the process of a node u activating one of his neighbours v with probability $p_{u,v}$, as flipping a biased coin and if it succeeds declare the edge *live*, otherwise declare it *blocked*. Moreover, we can w.l.o.g. consider that all the coins are tossed before the process begins. Therefore, from the initial graph $G(V, E)$, we get a graph $G(V, E_{live})$ where we keep only live edges. Now, in this setting all nodes that are reachable via a live path from the initial set A_0 would become active when the cascade process quiesced. This view is very helpful and will be used to prove a crucial property about our model. The model is also equivalent to the *bond percolation*[19] setting studied in probability and physics.

There are many generalization of the ICM. As a first step we can allow for the activation probabilities of individuals to be different between different pairs of nodes. Specifically, each active node u has a fixed probability $p_{u,v}$ to activate a non-active neighbour v. However, we still consider that the outcomes of each attempt are independent. The most general setting would be to allow the influence probabilities $p_{u,v}$ to depend on the subset S of v's neighbours that have already tried and failed. This fact would be encoded in a function $p_v(u, S) \in [0, 1]$ where $\{u\}, S$ are disjoint sets of neighbours of v. This model can result in different outcomes depending on the order of activations, so either we must define a specific order in which neighbours try to activate v or, even better, consider only functions p_v that are order independent. The class of functions that satisfy those constraints are those that for every set S that have tried and failed, for every set $X \subset S$, $p_v(u, X)$ depends only on the cardinality of X, i.e. the subset of S defines an incremental function on S. This model is called the *General Cascade Model*(GCM).

2.1.2 Linear Threshold Model(LTM)

The LTM stems from the early work of Schelling[123] and Granovetter[62]. It was used in a context of explaining collective behaviour on a non-normative basis; as a dynamic process of opinion formation. In this model, a node v is influenced by each active neighbour u according to a weight $w_{u,v}$, such that $\sum_{u\ v} w_{u,v} \leq 1$. We consider that every node is associated with a node specific threshold, i.e. amount of "influence" that is needed for him to change his mind. For this particular model the threshold is picked uniformly at random from the $[0, 1]$ interval and represents the weighted fraction of the node's neighbours that are needed for him to become active.

The process unfolds as follows. A set of initial nodes A_0 exerts some amount of influence to their neighbours. If the total influence that a node perceives is greater than the node specific threshold, then that node becomes active and in turn influences other nodes. The process continues until no other node is activated. The process proceeds in discrete steps: at each time step t, all nodes A_{t-1} that were active in step $t-1$ remain active, and all nodes in $V \setminus A_{t-1}$ that the total influence they perceive exceeds their individual thresholds become active:

$$A_t = A_{t-1} \cup \{v : \sum_{u \in A_{t-1},\ u \sim v} w_{u,v} \geq \theta_v\} \tag{2.1}$$

The number of steps T until the process settles is the first time that $A_t = A_{t-1}$: $T = inf\{t : A_t = A_{t-1}\}$. Obviously this means that $T \leq |V|$ for both processes.

Again there are many ways to generalize this model to incorporate diverse effects. One could argue that in that direction we must both allow for more general threshold functions

$f_v(S)$, instead of simple linear functions, and more general distributions F_v according to which the node specific thresholds are being selected, instead of the uniform distribution. It is easy to see that only the first requirement is needed, since any extra information provided by the threshold distribution can be incorporated in a more complex influence function f. The most general version of the model is the *General Threshold Model*. In this model we associate with each node v a monotone threshold function $f_v(S)$ that maps every subset of v's neighbours to a number in $[0, 1]$ with the condition that $f_v(\emptyset) = 0$. The process proceeds in the same way as the LTM with the only exception that nodes are activated when $f_v(S) \geq \theta_v$. The LTM is a special case for $f_v(S) = \sum_{u\in S,\ u\ \ v} w_{u,v}$, with parameters $w_{u,v}$ such that $\sum_{u\ \ v} w_{u,v} \leq 1$.

2.1.3 Model Equivalence

A natural question to ask, is how nuanced is the outcome of the process depending on which model we choose? Is the cascade or the threshold view more pertinent? Intuitively, both models cannot be substantially different. Take for instance the LTM, this model just asserts that if the influence that is exerted on a certain node exceeds a certain level, the node becomes active. On the other hand, in the ICM a single node has the chance to activate one of his neighbours. However, this happens with a probability proportional to the influence he has on that particular node and these influences are "added" in a probabilistic sense for different attempts. It turns out that both the ICM and LTM produce the same distribution over outcomes and are in that respect equivalent.

To prove the above statement we will proceed by proving equivalence under the live edge path viewpoint. We will accomplish that by induction. Consider that at time step t the set of active nodes is A_t. If a node v has not become active by the end of step t, then that means that the total influence perceived by that node is smaller than the chosen threshold, i.e.: $\sum_{u\in A_t} w_{u,v} \leq \theta_v$. But since θ_v is distributed uniformly in $[0, 1]$ then conditional on node v being inactive after step t, $theta_v$ is distributed uniformly in $(\sum_{u\in A_t} w_{u,v}, 1]$[1]. Therefore, the probability that a node becomes active in step $t + 1$ is:

$$\mathbb{P}[v \in A_{t+1}] = \frac{\sum_{u\in A_{t+1}} w_{u,v} - \sum_{u\in A_t} w_{u,v}}{1 - \sum_{u\in A_t} w_{u,v}} = \frac{\sum_{u\in A_{t+1}\setminus A_t} w_{u,v}}{1 - \sum_{u\in A_t} w_{u,v}} \tag{2.2}$$

To provide the connection with the live edge path view, we must decide on which edges are live and which edges are not. Actually, since the reachability relation is transitive relation it suffices to decide on whether a node is reachable by the active set A_t and not on the specific edges. Thus, at each time point t starting from $t = 0$ we decide for all the nodes that are adjacent to the set A_t whether their live edge comes from this set or not. So, a node v becomes reachable at time t with probability given by (2.2). By induction on the number of steps the set of active(reachable) nodes for the two models are equivalent.

The same reasoning can be applied to prove that the equivalence hold for the GTM and GCM as well. The only difference is in the way we convert between the probabilities $p_v(u, S)$ and the threshold functions $f_v(S)$. To convert between from the LTM to the ICM we must define a probability that a node u activates a neighbour v given that nodes in S have tried and failed. Again, if nodes in S have tried and failed then that means that θ_v belongs in $(f_v(S), 1]$.The probability that u activates v under these circumstances is :

$$p_v(u, S) = \frac{f_v(S \cup \{s\}) - f_v(S)}{1 - f_v(S)} \tag{2.3}$$

[1]This is due to the fact that the uniform distribution is self-similar at every (union) of sub-intervals.

Applying the same reasoning as before, it is easy to see that the cascade process is equivalent to the threshold process.

Towards the other direction, consider a node v and a set of active neighbours $S = \{u_1, \ldots, u_k\}$. The probability that a node is not activated is the joint probability that all neighbours independently tried and failed to activate v: $\prod_{i=1}^{k}(1-p_v(u_i, S))$. Because, we only consider function p_v that are order independent, we can define the corresponding threshold function f_v depending only on set S:

$$f_v(S) = 1 - \prod_{i=1}^{k}(1 - p_v(u_i, S)) \tag{2.4}$$

It is straightforward to show equivalence of the threshold process under these functionss to the original cascade process.

2.2 Submodularity

The crucial property that all our models satisfy is that of *submodularity.* We will see how this property can be translated into constant approximation algorithms for our problem of finding the best initial set A. Formally, a set function, i.e. function that takes as input subsets S of universal set U, is called submodular when:

$$f(S \cup \{v\}) - f(S) \geq f(T \cup \{v\}) - f(T) \tag{2.5}$$

for all $v \in U$ and all pairs of sets $S \subseteq T$. Intuitively, submodularity is the set-function analog of concavity. Specifically, a function is called submodular if it satisfies the "diminishing returns" property: the marginal gain by adding an element to a set S it is at least as the marginal gain by adding an element to the superset T. In other words, the higher the ground value is, the smaller is the marginal gain of adding one element.

The function that we are interested in is the *total influence function* $\sigma(A)$, that is the expected number of active nodes after the process terminates with an active initial set A. The expectation is taken with respect to the randomness of the model. In the live edge path view, this can be conveniently captured by specifying the set X of live edges. Therefore, the total influence function in the case of the ICM is:

$$\sigma(A) = \sum_{X} \mathbb{P}[X] \cdot \sigma_X(A) \tag{2.6}$$

We have managed to write the total influence function as a convex sum of influence functions where the outcome of the random coin tosses is fixed. So, if we can show submodularity of σ_X for all X, then it is a straightforward exercise to show submodularity of σ. To that end, we define $R(u, X)$ the set of all nodes reachable from u via a live edge path and since X is known when we are considering σ_X, $R(u, X)$ is a deterministic quantity. The outcome specific function $\sigma_X(A)$ can now be expressed as the cardinality of the set $\bigcup_{u \in A} R(u, X)$. To prove submodularity of σ_X consider two sets $S \subseteq T$ and the quantity $\Delta(S) = \sigma_X(S \cup \{v\}) - \sigma_X(\cdot)$, which is the number of elements in $R(v, X)$ that are not in the union $\bigcup_{u \in A} R(u, X)$. Since $R(v, X)$ is the same irrespectively of S or T and σ_X is monotone, then it follows that $\Delta(S) \geq \Delta(T)$, which is the defining inequality of submodularity. Thus, we have proven submodularity of the total influence function for the models that are equivalent in the live edge path viewpoint, which besides the ICM and the LTM includes some other generalizations as well, such as the

Triggering Model and *Decreasing Cascade Model* introduced in [76]. However, this method of proving submodularity breaks down when we consider the most general models GTM and GCM.

Results towards the latter direction were obtained in 2007 by Mossel and Roch[105], who proved the following conjecture of KKT[76].

Conjecture 1 (KKT)**.** *Consider the General Threshold Model, whenever the threshold functions at every node are monotone and submodular, the resulting influence function $\sigma(\cdot)$ is monotone and submodular as well*

This result is interesting in the sense that it shows that monotonicity and submodularity properties are closed under diffusion processes; that local monotonicity and submodularity is sufficient to induce these properties on a global aggregate level. Their proof uses carefully crafted coupling arguments for the stochastic process. We will only sketch the main ideas behind the proof. Initially, we require a different, but equivalent, definition of submodularity:

Definition 1. *The set-function $f : 2^V \to \mathbb{R}$ is submodular if for all $S, T \subseteq V$:*

$$f(S) + f(T) \geq f(S \cap T) + f(S \cup T) \tag{2.7}$$

In our case the total influence function σ plays the role of f and two arbitrary initially active sets A, B play the role of S, T. Let A_{n-1}, B_{n-1} denote the set of nodes that are active after the process terminates. It is easy to see that:

$$\sigma(A) + \sigma(B) = |A_{n-1}| + |B_{n-1}| = |A_{n-1} \cap B_{n-1}| + |A_{n-1} \cup B_{n-1}| \tag{2.8}$$

Note that we would be done if we could say that $|A_{n-1} \cap B_{n-1}| \geq \sigma(A \cap B)$ and $|A_{n-1} \cup B_{n-1}| \geq \sigma(A \cup B)$. The first condition is trivially satisfied by construction, since $\sigma(A \cap B)$ is a subset of both A_{n-1} and B_{n-1}. So, the real hurdle is to prove the second condition.

Intuitively, it is not clear at all that this is the case, as growing the two processes A_{n-1}, B_{n-1} separately does not give an advantage from a deterministic set-function perspective. Nevertheless, we can foresee a certain probabilistic advantage of considering separate processes. Imagine that some key node happens to choose a high threshold value and never quite gets to be activated, thus impeding the diffusion process. If we had both sets initiated in the same run, the extra potential could be wasted by random fluctuations of the threshold values. On the other hand, if we run the processes separately then such bad events would be somewhat ameliorated. This intuition was formalized and exploited by Mossel and Roch[105] in the following way:

- Instead of considering the sets $A, B, A \cap B, A \cup B$ as initial active sets, they choose another equivalent collection of sets $A \setminus B, B \setminus A, A \cap B$ which if combined produce the previous sets.
- They considered a process where the initially active nodes are introduced in bundles, where each bundle is introduced in the network after the process has coalesced for previous bundles. This was done in order to gain better control over the processes and distinguish between the effects of introducing different bundles. Specifically, all processes A_{n-1},B_{n-1},$(A \cup B)_{n-1}$ are grown in the following manner.

  1. The intersection bundle $(A \cap B)$ is introduced.
  2. The pure A bundle $(A \setminus B)$ is introduced(if exists).

3. The pure B bundle $(B \setminus A)$ is introduced(if exists).

So, we have that the A_{n-1} process is $(1, 2, -)$, the B_{n-1} process is $(1, -, 3)$ and the $(A \cup B)_{n-1}$ process is $(1, 2, 3)$.

- In order to exploit the fact that A_{n-1}, B_{n-1} processes are grown separately we must somehow stochastically couple the threshold values in a manner so that the union of the two processes is maximized. This was accomplished by considering that the threshold values θ_v of all vertices that were activated after the $\sigma(A)$ process(bundles 1 and 2) have the same distribution in the B process(bundles 1 and 3) and that the vertices that were not activated after the second bundle are distributed like $1 - \theta_v$, conditional that they were not activated by the A process($\theta_v \in (f_v(A_{n-1}), 1]$).

By these methods the authors were able to prove the necessary subset relations and then, by exploiting the local monotonicity and submodularity properties, inductively prove the desired inequality that we mentioned in the beginning. Summing up, the proof considered incremental growth of the processes, in order to use induction and exploit the local properties, and coupling to exploit the extra freedom of running the processes separately.

2.3 Greedy Approximation Algorithm

In general submodular functions have been extensively studied in the optimization literature [107, 49, 71, 46, 87, 54, 88, 130, 52, 10] and, thus, their properties are pretty well understood. Specifically, there is an old result of Nemhuaser et.al[107] that shows that the following greedy approximation algorithm approximates the optimal within a ratio of $(1 - 1/e)$ for *submodular* functions that are *monotone* and take *non-negativite* values. Therefore, if we prove these properties about our problem we would obtain a constant approximation algorithm.

Alg. 1 Greedy Approximation Algorithm

Input: $G = (V, E),\ k$
Initialize: $S = \emptyset$.
for k iterations **do**
1 find $u \in V \setminus S$ that maximizes $\sigma(S \cup \{u\}) - \sigma(S)$.
2 add u to S.
end
Output: Influence set A, with $|A| = k$

In the previous section, we showed that the influence function σ satisfies the submodularity property. Additionally, the monotonicity and non-negativity properties are trivially satisfied. However, since $\sigma(A)$ is actually the expected (weighted)number of active nodes by using an initial seed set A, it is not clear that *step 1* can be performed in polynomial time. If we were actually to calculate the exact expected value we would need to sum over all the expected outcomes, which for a graph with m edges would be of the order 2^m. Therefore, exact calculation is out of the question. What about approximate calculation? Is it possible to approximately compute the value $\sigma(A)$ within an error margin? The answer is affirmative. To see that, observe that a seed set A defines a distribution over $|A|, \ldots, |V|$ for the number of active nodes after the end of the process. Therefore, we might be able to estimate the expected value $\sigma(A)$, which is a deterministic function of A, by drawing independent samples from the distribution. Each sample can be extracted in polynomial time(e.g. biased coin flipping for the edges in case of the ICM), thus the real issue here is how many samples N are enough in order to get the desired approximation guarantee between the *empirical mean* $\hat{\sigma}(A) = \sum_{i=1}^{N} X_i$ and the actual expected value $\sigma(A)$. The following theorem clarifies the situation:

Theorem 1. (Chernoff-Hoeffding Bound[68]) *Let $X_1, \ldots, X_N$ be independent, identically distributed random variables with $0 \leq X_i \leq n$ and let $X = \sum_i X_i$, and $\mu = \mathbb{E}[X]$, then for all $\epsilon > 0$:*

$$\mathbb{P}[|X - \mu| \geq \epsilon\mu] \leq 2e^{-\frac{\epsilon^2 N}{n^2}}$$

Using, this theorem we can show that if we want to have an $(1 \pm \epsilon/2)$ estimate with probability at least $(1 - \frac{1}{n^2})$, we must use $N > \frac{n^2}{\epsilon} \log n$ samples. This implies that the correct value of the function $\sigma(A \cup \{u\})$ is estimated in worst case within $(1 \pm \epsilon)$ accuracy. Therefore, we have showed that *step 1* actually can be performed in polynomial time with arbitrary accuracy. What is left to do, is to actually show how the $(1 - 1/e)$ guarantee, results from subomdularity, non-negativity and monotonicity. The following theorem provides the answer:

Theorem 2. (Nemhuaser, Fisher, Wolsley[107]) *if f is a non-negative, monotone and submodular function, then the greedy algorithm is a $(1 - 1/e)$-approximation for the problem of maximizing $f(S)$ subject to the constraint that $|S| = k$.*

Proof. The proof of the theorem relies on the fact that submodularity and monotonicity guarantee us that when we choose a node to add to the set by maximizing the marginal gain, all future marginal contributions of other nodes will be at most equal. Let us formalize this statement. Assume that the greedy algorithm has selected the nodes in $S = \{v_1, \ldots, v_k\}$ and let S_i denote the nodes selected up to time i, that is $S_i = \{v_1, \ldots, v_i\}$. The marginal benefit from the addition of node i is $\delta_i = f(S_i) - f(S_{i-1})$. Now, assume that O is the optimal solution with k elements and let $\tilde{O}_i = O \cup S_i$ be the union of the two sets which has at most $k + i$ elements(when the two sets are disjoint). By monotonicity we trivially have that $f(O) \leq f(\tilde{O}_i)$ for all i. Note that the algorithm at each step i chooses the nodes with maximal δ_i, therefore in worst case the algorithm at step i would have selected i nodes that none of them belongs to the optimal set O. However, the fact that none node from set O was chosen means that all those nodes had always marginal contribution smaller from δ_t for $t = 1, \ldots, i$ and thus will have contribute to f at most $k \times \delta_{i+1}$ if we add them afterwards. So, we conclude that $f(O) \leq f(\tilde{O}_i) \leq f(S_i) + k\delta_{i+1}$. Also, note that since we construct set S_i incrementally we have that: $f(S_{i+1}) = f(S_i) + \delta_{i+1}$. Using the last two relations we get:

$$f(S_{i+1}) \geq f(S_i) + \frac{1}{k}[f(O) - f(S_i)] = (1 - \frac{1}{k})f(S_i) + \frac{1}{k}f(O) \tag{2.9}$$

which says that our greedy solution is "diluted" with at least $1/k$ fraction of the optimal solution at each step. Thus, we assume that $f(S_i) \geq (1 - (1 - \frac{1}{k})^i)f(O)$ and we use induction to prove it. The case $i = 0$ is trivial by non-negativity. The induction step unfolds as:

$$\begin{aligned}
f(S_{i+1}) &\geq (1 - \frac{1}{k}) \cdot f(S_i) + \frac{1}{k} \cdot f(O) \\
&\geq (1 - \frac{1}{k}) \cdot (1 - (1 - \frac{1}{k})^i) \cdot f(O) + \frac{1}{k} \cdot f(O) \text{(by induction hypothesis)} \\
&= (1 - (1 - \frac{1}{k})^{i+1}) \cdot f(O).
\end{aligned}$$

which completes the induction. It is an easy exercise to see that $(1 - \frac{1}{k})^i \geq \frac{1}{e}$ for all $i \leq k$. Thus, we get that our algorithm approximates the optimal within a ratio at least $(1 - \frac{1}{e})$. □

2.4 Computational Issues and Heuristics

The *Influence Maximization* problem, besides its theoretical appeal, is a practical one. Moreover, modern social networks such as Facebook or Tweeter boast hundred of millions of users. Therefore, although in theory the greedy approximation algorithm provides the best theoretical guarantee from a practical viewpoint it suffers because of extensive running time. Specifically, we might need roughly $O(nk)$ iterations of step 2 of the greedy algorithm and at least $\frac{n^2}{\epsilon^2}\log(n)$ samples to estimate the influence functions(in reality each step should require that much samples, however we can keep the outcome of a live-edge-path instance and use it to estimate all instances of $\sigma(S)$), which become prohibitive for large networks. So, we must work on two fronts: reducing the computational cost of estimating σ and also reducing the number of sets upon which we estimate σ. The hope is that we could salvage enough of the algorithm's performance inspite relaxing the computational strain.

The major advance towards the latter goal was achieved by Leskovec et al.[94]. The authors proposed in their paper, among other things, an algorithm called *Cost Effective Lazy Forward* Selection(CELF), where they exploit the submodularity property to greatly reduced the running time of the greedy algorithm. Specifically, when evaluating the marginal gain $\sigma(S \cup \{u\}) - \sigma(S)$, the authors store the information in a priority queue. Their main idea was that we don't need to re-evaluate the marginal gains for all vertices in $V \setminus S$, but only for those that are higher in the queue. When we find the first node that is larger than all the nodes in the queue(updated or not) we can safely stop updating, as submodularity guarantees us that even if we updated all the nodes their values would be at most equal but not greater. The authors experimentally evaluated the performance of the algorithm and showed that it drastically reduces the running time(up to 700 times).

A first improvement of the CELF algorithm[30] was to use UnionFind structures in order to calculate the spread of influence for nodes. Specifically, we perform initially R live-edge-path experiments and for each one via BFS process create UnionFind structures. Therefore, when we are asked to calculate the influence of a single node is just the cardinality of the Union(strongly connected component) it belongs to thus we can calculate the values $\sigma(\{v\})$ while we are performing the experiments in almost linear time. There is also a recent improvement of this algorithm called CELF++[61], where besides the marginal gain $\delta_u(S)$ the algorithm uses the same Monte Carlo simulations to estimate additionally $\delta_u(S \cup \{current_best\})$. Experiments indicate that this algorithm has a almost negligible memory overhead than CELF but improves running time by $30 - 50$ percent.

These algorithms don't sacrifice the strong $(1-1/e)$-approximation guarantee and manage to improve the scalability of the problem. Nevertheless, they still are unable to handle very large networks with million of nodes as there is an inherent quadratic term in performing the Monte Carlo estimations of the influence function σ, which cannot be dropped unless we are ready to forfeit the theoretical approximation guarantees. In that direction, there are many efforts to approximate the influence function.

Kimura and Saito[80] aiming to overcome the computational burden of evaluating the influence function proposed two models SP and SP1 that "approximate" in a certain sense the IC model. The first model considers that a node can be activated from a set A only via shortest paths, whereas the second model allows nodes to be activated again only by shortest paths but additionally only at times $d, d+1$ where d is the hop-distance between the node and the seed set A. The intuition behind this proposal is that along a path between a seed set and a node, the probability that a node is activated by the seed set is maximal along the shortest path. Therefore, by only considering those cases we might get a good approximation. Moreover, we expect that when as the average influence probabilities gets smaller and smaller this approximation should get better. The authors performed computational experiments where they also considered the ranking similarity by running the greedy algorithms on these models and the original IC model. The results showed that indeed the ranking of nodes provided by these models are comparable to those obtained by the IC model and better than some heuristics based on degree or eigenvector centrality, especially when the influence probabilities are small.

Another approach on heuristically obtaining influential nodes was proposed by Chen et al.[30]. They introduced a heuristic called *DegreeDiscount*, where the main idea is to incrementally construct the influence set A by adding high degree nodes but making sure to update the degrees of the reamining nodes by removing the edges that reside in the current seed set S. When the influence probabilities are small the authors modified this heuristic to approximate the influence in the IC model. Additionally, they conducted computational experiments which showed that in many cases this heuristic performs reasonably well and in time that is greatly smaller($O(k \log n + m)$ compared to $O(knNm)$ of the original greedy algorithm).

Despite the advances, researchers were not satisfied neither with the running time of the SP model nor with the performance of the DegreeDiscount heuristic and continued to search for better methods for finding good seed sets. A major advance was made by Chen, Wang and Wang[29] with the PMIA(Prefix Excluded Maximum Influence Arborescence) algorithm. The authors took the SP idea of Kimura and Saito one step further. Instead of considering shortest paths(in the hop-distance sense) they consider *maximum influence paths* and to reduce running time the consider only one unique path between two nodes. Therefore, the computation of the influence spread is reduced to finding for each node an arborescence(directed rooted tree) which can be done efficiently by Dijkstra like algorithms. The authors manage to reduce the calculation of global influence spread in terms of the local influence regions(arborescences) of nodes and control the size of those regions by a parameter as a trade-off between quality and speed. The algorithm performed on par or better than the SP1 model and with significantly improved running time, while it outperformed previous heuristics such as DegreeDiscount and PageRank. In a related paper[31], Chen,Yang and Zhang apply these ideas in the case of the Linear Threshold Model as well.

One drawback of the PMIA approach is the extensive memory usage that is associated with maintaining the local tree structures, which can reach tens of GB's for large graphs and slow down the overall running time. Jung,Heo,Chen[73] considered a different approach integrating the MIA and the PageRank idea for the IC model. They first proposed a method of estimating the individual influences $\sigma(\{u\})$ to select the first node, by noting that in trees the problem has a nice recursive formulation that results in a system of linear equations and then generalizing this approach in a message passing algorithm that provides an global influence ranking. Then the authors overcame the problem of the inherent overlap in influence of ranking methods, by using local tree structures to approximate the marginal influence of individual nodes given a current seed set. They called their algorithm Influence Ranking and Influence Estimation(IRIE). Experimental results showed that IRIE performs as good as PMIA while requiring a small fraction of the memory usage and being up to two orders of magnitude faster. Another significant advantage is that IRIE's simple iterative computation can be implemented in a parallel graph computation platform, thus promising further speed up and scalability.

Lastly, there is also a recent work of Goyal et al.[60] that considers heuristics for the LTM model, improving on work of [31]. The author address some limitations of the LDAG approach, namely the increased memory usage and the hardness of finding a good LDAG. Using a decomposition of the total influence as a sum of influences of individual nodes in the graph induced by removing all seed nodes but one, which was first shown in [129], they reduce global influence computation to local. Particularly, they address the issue of hardness of finding a good LDAG by considering paths only in a k-neighbourhood and thus improving influence estimation running time. Moreover, since influence of a node can be expressed via the influence of it's neighbours, the authors reduce the total number of influence estimation calls by considering only nodes in a heuristically computed Vertex Cover of the graph. Finally, they considered a look-ahead rule in order to evaluate the marginal influences and called their algorithm Simpath. Experiments conducted in real social network graphs showed that Simpath significantly improves Ldag's running time and memory usage while achieving a slight improvement on the solution quality.

2.5 Further Reading

The *Influence Maximization(IM)* problem is of great importance in our contemporary world, were social networks are ubiquitous and users spend increased amount of their time online. Companies are, thus, motivated to actually implement and apply the ideas developed over the years by the IM community. However, there are some caveats in doing so.

First and foremost, the IM setting assumes that we know both the social graph as well as the influence probabilities, which in reality could not be further from truth. There is a line of research involved with finding ways to extract graphs and infer influences from online observations. Saito et al.[121] considered the problem of inferring the influence probabilities from a set of activation observations $D(0), D(1), \ldots, D(T)$, where $D(t)$ denotes the set of nodes that were activated at time t. They used a Maximum Likelihood criterion and solved it using an Expectation-Maximization(EM) approach. They also extend their methods when considering also asynchronous time delay episodes[120]. On a related work, Goyal et al.[59] considered different ways of relating action log(events) and influence probabilities under various models (static,continuous,discrete). They designed efficient algorithms for learning those parameters and experimentally evaluated their performance. There is also work of Leskovec and co-authors that deal with these issues and is similar in nature to the above approaches[106, 117, 58]. An interesting direction is also that of predicting links and probabilities from prior heterogenous data[69, 79]. Finally, there is also an attempt to design mechanisms were the planner can extract the influence probabilities from the buyers themselves by considering appropriate payments of a VCG-mechanism[103] or to extract individual costs associated with influencing particular individuals[125].

Moreover, the context of IM concerns binary non-progressive decisions. That is, a buyer can only change his mind once. Reality is much more complex. In real life sometimes multiple products compete to satisfy particular needs of buyers and sometimes not only positive information diffuses through a network. Hence, there are efforts to simultaneously incorporate diverse phenomena in the diffusion process and many extensions of the basic models have been proposed.

First, KKT[76] managed to reduce the *progressive* case, were buyers can change their minds, to the *non-progressive* classic setting by considering layered graphs. Additionally, the authors in the same paper considered more general marketing strategies were there are more than one actions available and the payoff depends continuously on the amount invested. The issue of competition and compatibility between two products was discussed in[70, 26, 14, 4, 22] and most results are in spirit of the original KKT paper. Also, work of Chen et al.[28] covered the case whether both negative and positive opinions about a product can propagate through the network. They proved that the greedy algorithm also works in this setting and showed that the process is sensitive to the negativity bias(probability of product dissatisfaction) via a Price of Anarachy[85] approach.

The case of multiple products was first considered in[112]. The authors introduced a model where k products compete and the final state of the graph S is sampled from the ensemble of possible states $(k+1)^{|V|}$ by a rapidly mixing Markov Chain. The distribution defined on those states is derived by the aggregation of a local rule were both an arbitrary switch of a node and a biased switch based on the colors of it's local neighbourhood is allowed. Recently, Markakis and Apt[6] generalized the LTM for allowing multiple products and provided answers to questions such as: when is the outcome of the process unique or when does a product dominate the whole network. They also consider the computational hardness of some related decision problems.

Another aspect of the problem is whether the formulation and assumptions made by the IM community are relevant to real world situations. So, there are many empirical studies looking into the way information and influences cascade through networks, blogs, etc. This line of research[82, 96, 9, 90, 95, 98, 84, 38, 91, 92, 36, 116, 89, 118, 32, 126] also aims at obtaining a more qualitative anthropomorphic understanding of the process and rough guidelines for managers in order to be able to design effective policies.

Lastly, for more information related to information cascades and networked phenomena, there is the excellent book of D.Easly and J.Kleinberg[39] and reviews from Kempe[135] and Wortman[136].

Chapter 3

Revenue Maximization

The importance of social networks in advertising and market research is by now indubitable[44, 114, 131, 21]. Social networks provide us with detailed and broad information for millions of users and companies have been using this information to increase market penetration of their products. Social network companies use the revenue exerted by advertisements to sustain the costs involved in maintaining their servers and quality of service, as well as to provide the basis of a sustainable business model. However, there is a large discrepancy between the perceived value of Social Networks and the actual revenue they generate. For instance Facebook, a 7 year old company, was valued recently by Goldman Sachs at 50 billion dollars but actual revenues were estimated to be around 2.2 billion dollars(eMarketer). The widespread belief is that much of the potential of Social Networks remains unexploited.

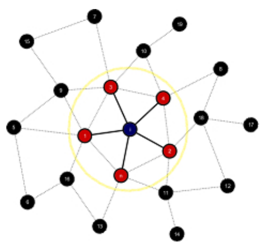
This premise has spurred a large amount of research in the direction of monetizing Social Networks. The current business model is organized around the sponsored search paradigm of contextual advertising, which, though successful, disregards at large the network effects between buyers. Researchers have acknowledged the importance of network structure and have began to study the impact of ,the so called, *network externalities* in a variety of settings. We are interested in how can a seller exploit these externalities to design intelligent marketing strategies to maximize his revenue. In our context there are two complementary ways that externalities come into play.

Propagation. Information about products diffuses through the social network via personal communications. Consumers share their opinions about products and often are motivated to suggest an exciting new product or condemn a dysfunctional–disappointing purchase. Often these recommendations are far stronger than advertising attempts and, whether they are true or misleading, have great impact in the final adoption of the product.

Network Value. There are settings where the utility of the product depends inherently on the scale of adoption that the product receives. The Internet, Windows OS or Facebook are examples of "products" whose value lies in the fact that a large fraction of the population has access and uses them on a daily basis. Moreover, for many products (e.g. cell phones, instant messengers, online gaming) the value of the product for a buyer depends on the specific set of his friends that have adopted it and possibly in a non-uniform way.

In this and the following chapters we focus on the second kind of externalities, were product adoption is an integral part of product value. In this setting, the seller must devise a marketing strategy that guarantees both wide product adoption(increased value of the product or service) and significant revenue. We follow up on a work by Hartline, Mirrokni, Sundararajan(HMS)[66]. They considered the case of a single seller selling a single product to a set of potential buyers under *positive network externalities*, where buyers are willing to purchase the product in a higher price if some of their friends already bought it.

3.1 Influence Models



Following the Influence Maximization paradigm [76], HMS model buyers and their relationships by a graph. The world consists of n buyers that have arbitrary social relationships between them, these relations influence buyers in their decisions. Buyers are modelled by *nodes* and the influence between them by weights w_{ij}, $\forall i, j \in V$. Specifically, every buyer has a valuation $v_i(S)$ for the product, which depends at any time point only from the set of buyers S that already own the product. The exact quantities v_i for the good are unknown to the seller and are treated as *random variables* of which only the distributions $F_{i,S}$ are known for all $S \subseteq V$ and for all $i \in V$. The values $v_i(S)$ are assumed to be independently distributed.

The above model is quite general and encompasses the inherent uncertainty of buyers preferences by treating the valuations as random variables. Some important instantiations of the general model are:

Uniform Additive Model In this model there are deterministic weights $w_{i,j}$ for all $i, j \in V$. Given the set of buyers S that already own the product and a asking price x, buyer i decides to accept the offer or not by adding all the weights from his active neighbours $\sum_{j \in S} w_{j,i}$ and then choosing a threshold in $\theta_i \in [0, 1]$. If the asking price x is greater than $\theta \cdot \sum_{j \in S} w_{j,i}$ then he rejects the offer, else he accepts and pays the asking price x. This model is a direct extension of the Linear Threshold Model studied in the context of Influence Maximization in order to incorporate pricing dynamics. Also note that we could equivalently say that buyer i chooses a random number in $[0, \sum_{j \in S} w_{j,i}]$, however the threshold interpretation is a more useful viewpoint as we will see in the following chapters.

Symmetric Model This is the simplest possible model, where buyers valuations $v_i(S)$ are identical and depend only on the number of buyers that already own the product $|S| = k$. Specifically, the model is fully specified by the distribution F_k for $k = 0, \ldots, n-1$. Note that this is an aggregate model of interaction where the identities of buyers play no role, i.e. there is no underlying graph. This model is reminiscent of the Baas model and the general "mean field" assumptions made in the Diffusion of Innovations context.

Concave Graph Model In this model, there are random weights $w_{i,j}$ for every buyer $i, j \in V$ each drawn independently from a distribution $F_{i,j}$. Each buyer is associated with a non-negative, monotone and concave function $f_i : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$. The value $v_i(S)$ is equal to $f_i(\sum_{j \in S \cup \{i\}} w_{j,i})$. In other words, we are uncertain about how buyers influence each other but we assume that the valuations depend only on the weighted aggregate of the influence and in

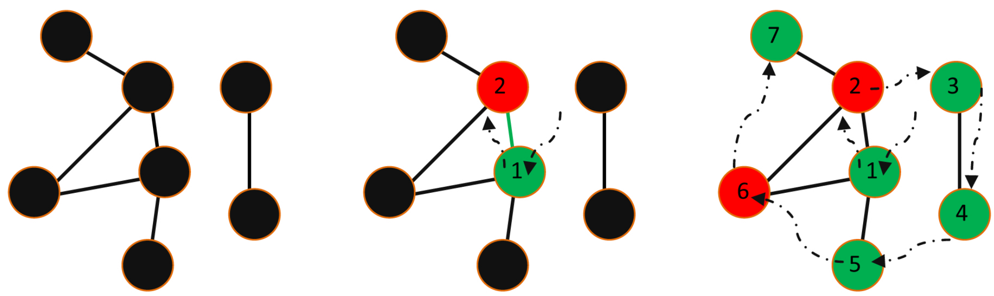


Figure 3.1: Illustration of a run of marketing strategy. Initially all buyers are available to receive an offer. The seller selects an ordering of buyers and makes individual offers. The buyers can either accept(green) or reject(red) the offer. If they accept they influence(green arcs) other buyers that have not received an offer yet.

a monotone, concave way. This model is analogous to the General Threshold Model were the threshold functions were monotone and submodular.

3.2 Optimal Marketing Strategies

In this setting Hartline et.al[66] search for strategies that maximize the revenue from a marketing strategy. They assume that the seller has the freedom to make individualized offers to buyers, that the cost of manufacturing a unit of the good is zero and that the seller has unlimited supply(e.g. software, electronic subscriptions). A *marketing strategy* consists of two elements: a sequence, according to which the seller approaches buyers, and the prices offered to each buyer. At each time point a buyer is visited and an offered is made. The buyer on his part can either accept the offer by paying the asking price or reject it. The revenue that results from a strategy is the sum of the prices paid by buyers that accepted the offer. We say that a marketing strategy is optimal when it maximizes the revenue.

Revenue Maximization The problem of *Revenue Maximization* is given a weighted graph $G(V, E, w)$ and an underlying influence model to find the marketing strategy(sequence of offers and price for each offer) that maximizes the expected revenue.

Hardness Hartline et.al. showed that this problem is *NP-Hard* for the directed case even for a very simple model, under complete information and linear valuation functions, utilizing a reduction from *Maximum Acyclic Subgraph Problem*[12, 67, 108]. Specifically, they considered a special case of the concave graph model were all $F_{i,j}$ are just degenerate point distributions($w_{i,j}$ are known) and the functions f_i are all the identity function $f(x) = x$. Under, these assumptions the seller needs only to find the right sequence π of offers to be made so as to maximize the revenue, which is just the sum of all the edges $w_{i,j}$ such that $\pi(i) < \pi(j)$. The problem of finding a sequence such that the number of directed edges going backwards is known as Maximum Feedback Arc Set and is known to be NP-Hard. Therefore, if for the directed case our problem is NP-Hard for this simple model, it is also for more general models. We must note at this point that the NP-Hardness for the undirected case does not follow from the reduction. Nevertheless, it was believed that the problem must be hard even for the undirected case, however no proof was known. We will address this issue in the following chapter.

Interestingly, a random permutation provides an 0.5 approximation for the Maximum Feedback Arc set, since the event $\pi(i) < \pi(j)$ happens exactly with probability 0.5 for all $i \neq j \in V$. Recently, Guruswami et al.[64] showed that assuming the Unique Games Conjecture it is NP-Hard to approximate Maximum Feedback Arc set by a factor greater than 0.5. This result implies that the random ordering is actually the best we can do. In the next chapter we will revisit this fact.

Dynamic Programming The previous Hardness result renders any effort of finding an optimal marketing strategy vain, at least in the directed setting. It seems like that when allowing buyers to have arbitrary relationships the problem becomes hard. What happens if get rid of the combinatorial dependencies between buyers and consider only aggregate phenomena? HMS looked into the special case of the Symmetric Model and showed that we can find an optimal solution via Dynamic Programming.

In the Symmetric Model, the sequence of buyers play no role(all buyers are identical). They assume that at each time step we offer our product to one of the remaining buyers at a price p. The price $p(k,t)$ only depends on the number of buyers k that have accepted the offer, and the number of remaining buyers t. Let $R(k,t)$ be the maximum revenue that can be extracted, when k buyers have accepted the offer and there are t more buyers to consider. Our strategy will be to construct a recurrence for the optimal revenue $R(0,n)$ in terms of the variables $R(k,t)$ and $p(k,t)$. They also assume that we know the distributions F_k and $f_k = \frac{dF_k}{dp}$.

We calculate the expected revenue from an offer of value $p(k,t)$. If the buyer accepts, which happens with probability $1 - F_k(p)$, then we get p plus the future optimal revenue $R(k+1,t-1)$ from the remaining buyers. Otherwise, we get only the revenue $R(k,t-1)$. The expected revenue is thus:

$$R(k,t) = F_k(p) \cdot R(k,t-1) + (1 - F_k(p)) \cdot [R(k+1,t-1) + p] \tag{3.1}$$

So, given that we have recursively computed $R(k,t-1)$ and $R(k+1,t-1)$ we can find the optimal price $p(k,t)$ by optimizing the expected revenue:

$$f_k(p)[R(k,t-1) - R(k+1,t-1) - p] + 1 - F_k(p) = 0 \tag{3.2}$$

Then the value $R(k,t)$ is easy to compute. To solve the recurrence we roughly need to fill an $n \times n$ matrix, thus we need roughly time $O(n^2)$. Therefore, in the symmetric model the optimal marketing strategy can be found in polynomial time.

The authors also investigated the special case were F_k is uniform in $[0, k+1]$. They solved the recurrence and found that the optimal strategy initially offered to a significant fraction of buyers the product for free(at zero price) and then the price offered to each successive buyer increased linearly. It is quite interesting that the optimal strategy involves two stages one "seeding phase", were the network value of the product is built up, and an "reaping phase", were the price increases gradually with time.

3.3 Influence and Exploit

The established hardness results bring about the need for approximation algorithms. Hartline et al., motivated by the fully symmetric setting, proposed a class of approximation strategies called *Influence and Exploit*(IE). These strategies consist of two steps. They initially offer the product for free to a selected subset A of buyers, aiming to increase the network value of the product(*Influence Step*) and then visit the remaining buyers in a random sequence extracting the maximal *myopic* revenue from each one(*Exploit Step*). In the fifth chapter, we thoroughly discuss the motivation for considering these kind of strategies, as well as provide our own improvements and insights. Here, we will just outline the approach taken by HMS. In general, IE strategies differ in how they construct the influence set A and how the revenue from this construction relates to the optimal revenue depending on the model we are considering.

Uniform Additive Model In this important case, when a buyer accepts an offer, the revenue we collect depends on the edges that are active at that time. Therefore, the Influence step aims at activating vertices such that a large fraction of the edges have exactly one active end(node). The authors proposed to construct the set A by including every vertex with probability q. The exploit step on the other hand requires just to visit the remaining buyers in a random order and offer the product at the myopic price.

The myopic price is the price which maximizes the expected revenue collected from the buyer, disregarding his influence in the network, i.e. how much more would a neighbour be willing to pay(in expectation) if the buyer we are considering accepted the offer. Let S be the set of buyers that have accepted the offer when we are considering buyer i, then the expected revenue from i is:

$$R_i(x) = (1 - U(x)) \sum_{j \in S} w_{j,i} = (1 - \frac{x}{\sum_{j \in S} w_{j,i}}) \sum_{j \in S} w_{j,i} \tag{3.3}$$

Setting $M_{i,S} = \sum_{j \in S} w_{j,i}$ and optimizing R_i with respect to x, we get that the optimal price is $M_{i,S}/2$. This price corresponds to $1/2$ probability that buyer i accepts. Hence, the myopic price in the case of the UAM is simply a price $x(i, S = M_{i,S}/2)$ such that buyer i accepts with probability $1/2$.

To analyze the performance of this algorithm, we need upper bounds on the optimal revenue. An obvious choice is to consider that the optimal strategy manages to activate all edges and extract the maximum amount(myopic) of revenue from each one. Thus, an upper bound for the undirected case would be:

$$R^*_{un} = \frac{1}{4} \sum_{(i,j) \in E} w_{i,j} == \frac{1}{4}(N + W) \tag{3.4}$$

where $N = \sum_{i \in V} w_{ii}$ and $W = \sum_{i > j} w_{ij}$ and $1/4$ is a factor results from the myopic revenue($(1/2) \cdot (1/2)$). Respectively this bound for the directed case would be:

$$R^*_{dir} = \frac{1}{4}(\sum_{i \in V} w_{ii} + \sum_{i > j} \max\{w_{ij}, w_{ji}\} \tag{3.5}$$

Now, every vertex i is included in the influence set A with probability $\mathbb{P}(i \in A) = q$ and in the exploit set E with probability $\mathbb{P}(i \in E) = 1 - q$. Moreover, every buyer in the exploit set accepts with probability $1/2$ and pays $M_{i,S}/2$, where S is the set of buyers that have accepted the offer. Recall that we are visiting buyers in the explout step in a random order, hence

$P(\pi j < \pi(i)) = 1/2$. Therefore, the expected revenue is:

$$\begin{aligned} R(q,1/2) &= \sum_{i\in V} P(i\in E)P(i)\left[\sum_{j\neq i} w_{ji}[P(j\in A)+P(j\in E\wedge j<i\wedge j\in S)]+w_{ii}\right] \\ &= \frac{1-q}{4}\left[\frac{1}{2}\sum_{i\neq j} w_{ji}(q+\frac{1-q}{4})+\sum_i w_{ii}\right] = \\ &= \frac{1}{4}\left[(1-q)N+\frac{1+2q-3q^2}{8}E\right] \end{aligned}$$

Now, we can optimize the expected revenue with respect to the parameter q. By differentiating and equating to zero we get that the optimal value is $q^* = \frac{W-N}{3W}$. The approximation ratio thus depends only on the ratio $\lambda = N/W$ and the worst case is when $\lambda = 0$, in which case $q = 1/3$. The approximation ratio is thus at least $2/3$.

Monotone Hazard Rate The approach taken in the case of the uniform additive model, can be generalized for the concave graph model in the case that the distributions F_{ij} satisfy the so-called monotone hazard rate condition.

Definition 2. *The hazard rate h of a distribution with density function f, cumulative F and support $[a,b]$ is $h(t) = \frac{f(t)}{1-F(t)}$. This implies that the distribution function can be expressed in term of the hazard rate $F(t) = 1 - \exp^{-\int_a^t h(x)dx}$.*

A function f satisfies the monotone hazard rate condition if the corresponding hazard rate function h is monotone and non-decreasing. This condition is roughly equivalent to the fact that when considering the expected revenue from a buyer there is a unique maximizer. The existence of a unique maximizer allows for a clearly defined price for each buyer in the exploit step irrespectively of the outcome of previous offers and the sequence of buyers.

Before, presenting the generalized analysis of the previous IE scheme, we state some properties about distributions satisfying the monotone hazard rate condition.

Lemma 1 (HMS[66]). *Let W_{ji} be random variables with distributions F_{ji} satisfying the monotone hazard rate condition. Let $Y = \sum_{j\in S} W_{ji}$ the random variable with distributionF_S. Also, let $R(x)$ be a monotone function and Z is a random variable with distribution $F_{i,S}$ derived from $Z = R(\sum_{j\in S} w_{ji})$. Then:*

1. *The distribution F_S of the random variable $Y = \sum_{j\in S} W_{ji}$ satisfies the monotone hazard rate condition.*
2. *The distribution $F_{i,S}$ of Z satisfies the monotone hazard rate condition for all $S \subset V$.*
3. *If a random variable X satisfies the monotone hazard rate condition then $1-F(x^*) \geq e^{-1}$, where x^* is the value that maximizes the function $(1-F(x))\cdot x$.*

In other words, points 1 and 2 say that the monotone hazard rate property is closed under summation and under application of a monotone function. In the concave graph model case this implies that the distribution of buyers valuations satisfy the monotone hazard rate condition. The third point just provides a lower bound on the probability that a buyer accepts an offer that maximizes the expected revenue.

To extend the simple IE strategy presented for the Uniform Additive Model for the special case of the Concave Graph Model where the weight distributions satisfy the monotone hazard rate condition, we need another lemma.

Lemma 2 (HMS[66]). *Consider a monotone submodular function $f : 2^V \to R$ and a subset $S \subset V$. Consider a random set $S^{'}$ by choosing each element of S independently with probability at least q. Then $\mathbb{E}[f(S^{'})] \geq q \cdot f(S)$.*

Proof. We will present the proof of the lemma as it is quite elegant. The idea is to express the revenue of the final set $S^{'}$ as incremental improvements from the empty set by telescopic sums. Now, for the terms of the sums to cancel each other alternatingly after taking expectation over the random choices, we need every random binary choice whether to include a vertex or not to have the same bias q. So, fix a permutation of vertices in S and consider that S_i is the resulting set after the first i vertices in the permutation have been considered. We have that $f(S^{'}) = \sum_{1 \leq i \leq |S^{'}|} f(S_i^{'}) - f(S_{i-1}^{'})$ and that $f(S_0^{'}) = 0$.Taking expectation, with respect to the set $S^{'}$, we get:

$$\begin{aligned}
\mathbb{E}[f(S^{'}] &= \mathbb{E}[\sum_{1 \leq i \leq |S^{'}|} f(S_i^{'}) - f(S_{i-1}^{'})] \\
&= \sum_{1 \leq i \leq |S|} q \cdot [f(S_{i-1}^{'} \cup \{u_i\}) - f(S_{i-1}^{'})] \\
&\geq \sum_{1 \leq i \leq |S|} q \cdot [f(S_i) - f(S_{i-1})] \\
&= q \cdot f(S)
\end{aligned}$$

where the inequality follows from the fact that $S_i^{'} \subseteq S_i$ for all i and that f is a submodular function. □

Now, we are ready to state the theorem:

Theorem 3 (HMS[66]). *Suppose that the revenue functions $R_i(S) = f_i(\sum_{j \in S} W_{ji})$ for all $i \in V$ and $S \subseteq V$ $\{i\}$ are monotone, non-negative and submodular and the distributions $F_{i,S}$ for all $i \in V$ and $S \subseteq V$ $\{i\}$ satisfy the monotone hazard rate condition. Then there exists a set A for which the simple IE strategy is a $\frac{e}{4e-2}$-approximation of the optimal marketing strategy.*

Again the IE strategy constructs the influence set A by including every vertex with probability q and then offers the myopic price at each vertex in $V \setminus A$ in a random order. The revenue is : $R(q) = \mathbb{E}_{A,T_i}[\sum_{i \in V \setminus A} R_i(T_i)]$, where T_i is the set of buyers that own the product when i is considered. The above expectation is with respect to the set A, the random sequence and buyers decisions. Every vertex in A is surely in T_i and every vertex in $V \setminus (A \cup \{i\})$ is in T_i with probability greater than $\frac{1}{2e}$, since there is an $1/2$ chance of being considred before i and, if so, there is at least a $1/e$ probability of accepting the offer. Therefore, for buyer $i \in V \setminus A$ every other vertex is in T_i with probability at least $(q + \frac{1-q}{2e})$. Also, every buyer is in $V \setminus A$ with probability $(1 - p)$. Thus, using the previous lemma we get that:

$$R(q) = \mathbb{E}_{A,T_i}[\sum_{i \in V \setminus A} R_i(T_i)] \geq (1-q)(q + \frac{1-q}{2e}) \sum_{i \in V} R_i(V \setminus \{i\}) \tag{3.6}$$

Note that in our case an upper bound on the revenue is $\sum_{i \in V} R_i(V \setminus \{i\})$, hence the approximation ratio is $(1-q)(q + \frac{1-q}{2e})$ which by choosing q appropriately is at least $\frac{e}{4e-2}$.

Submodular Valuations Hartline et al. aiming to extend their IE approach for more general valuation functions went to greater lengths. They considered the case where the only constraint on buyers valuations $R_i(A)$ is that they are non-negative and submodular, which are the weakest possible assumptions usually made in such context. Maximizing a non-monotone submodular functions is known to be NP-Hard[46] and therefore the authors turn to approximation strategies.

Observe that in this setting we do not hope to approximate the optimal strategy but instead we are aiming to approximate the optimal IE strategy. Let $g(A)$ be the expected revenue, with respect to buyers decisions and the random ordering, of an IE strategy with influence set A. Let O denote the optimal set, that is $g(O) \geq g(S)$ for all $S \subset V$. We are looking for a set A such that resulting IE revenue $g(A)$ is close to the optimal one.

The way HMS went about it was to exploit the local optimality conditions of submodular functions to iteratively construct a "good" set A. We first state some facts about non-negative submodular functions.

Definition 3. *Given $f : X \to R$, a set S is called a $(1+\alpha)$-approximate local optimum, if $(1+\alpha)f(S) \geq f(S \setminus \{u\})$ for any $u \in S$ and $(1+\alpha)f(S) \geq f(S \cup \{u\})$ for any $u \in X \setminus S$.*

This is a relaxation of the local optimality condition satisfied by submodular functions which says that if S is a local optimum then $f(I) \leq f(S)$ for every $I \subseteq S$ and for every $I \supseteq S$. However, this relaxation is necessary because finding an exact optimum of Max-Cut, special case of submodular function, is PLS-Complete[111]. We now present a Local Search algorithm to incrementally construct the influence set A.

Local Search

Input: *The set V and a value oracle for $g(S)$ on support 2^V*
Initialization:
1. *Set $S := \{v\}$ for the singleton set $\{v\}$ with the maximum value $g(\{v\})$ among singletons.*
2. If *there exists an element $v \in V \setminus S$ such that $f(S \cup \{v\}) > (1 + \frac{\epsilon}{n^2})f(S)$,* ***then*** *let $S := S \cup \{v\}$, and* ***go back*** *to Step 2.*
3. If *there exists an element $v \in S$ such that $f(S \setminus \{v\}) > (1 + \frac{\epsilon}{n^2})f(S)$,* ***then*** *let $S := S \setminus \{v\}$, and* ***go back*** *to Step 2.*
Output: *the maximum of $f(S)$ and $f(V \setminus S)$.*

The algorithm is constructed in a way that, when it terminates, produces an approximate local optimum. To analyze the performance of the algorithm we need also the following lemma.

Lemma 3 (FMV[46])**.** *If S is an $(1+\alpha)$ approximate local optimum for a submodular function f then for any set T such that $T \subseteq S$ or $T \supseteq S$, we have that $f(T) \leq (1 + n \cdot \alpha)f(S)$, where $n = |V|$.*

Proof. We will prove the lemma for the case $T \subseteq S$ and the other case is similar. Since $T \subseteq S$, we start from T and add vertices in $S \setminus T$ one by one $T_1 = T \subseteq T_2 \subseteq \ldots \subseteq T_k = S$ such that $T_i \setminus T_{i-1} = \{v_i\}$. By submodularity we know that $f(T_i) - f(T_{i-1}) \geq f(S) - f_(S \setminus \{v_i\})$ and since S in approximate local optimum we get that: $f(T_i) - f(T_{i-1}) \geq -\alpha f(S)$. The difference form a telescopic sum and if we sum all inequalities we get that: $f(S) - f(T) \geq -k\alpha f(S)$ and thus $f(T) \leq (1 + k\alpha)f(S) \leq (1 + n\alpha)f(S)$. ☐

We are now ready to state the theorem:

Theorem 4 (FMV[46]). *The Local Search is a $(\frac{1}{3}-\frac{\epsilon}{n})$-approximation algorithm for maximizing nonnegative submodular functions. The algorithm uses at most $O(\frac{1}{\epsilon}n^3 \log n)$ oracle calls.*

Proof. Consider an optimal solution O and fix $\alpha = \frac{\epsilon}{n^2}$ for some $\epsilon > 0$. If the algorithm terminates S is a local optimum an therefore $(1+n\alpha)f(S) \geq f(S \cap O)$ and $(1+n\alpha)f(S) \geq f(S \cup O)$. Using submodularity we get that $f(S \cup O) + f(V \setminus S) \geq f(O \setminus S) + f(V) geq f(O \setminus S)$ and $f(S \cap O) + f(C \setminus S) \geq f(C) + f(\emptyset) geq f(O \setminus S)$. If we combine the above inequalities we get:

$$\begin{aligned} 2(1+n\alpha)f(S) + f(V \setminus S) &\geq f(S \cap O) + f(S \cup O) + f(V \setminus S) \\ &\geq f(S \cap O) + f(O \setminus S) \geq f(O) \end{aligned}$$

So, if $f(S) < (\frac{1}{3} - \frac{\epsilon}{n})f(O)$ then $f(V\S) \geq (\frac{1}{3} - \frac{\epsilon}{n})f(O)$ and the algorithm outputs $f(V \setminus S)$, otherwise it outputs $f(S)$.

Our algorithm after each iteration improves the objective value by at least $(1 + \epsilon/n^2)$. Moreover, since f is submodular(diminishing returns property) it is easy to see that $f(S) \leq f(O) \leq nf(\{v\})$, where $\{v\}$ is the set returned after the first iteration. Therefore after k iteration we have that $f(S) \geq (1 + \frac{\epsilon}{n^2}{}^k f(\{v\})$. Hence, $k = O(\frac{1}{\epsilon}n^2 \log n$ and the total number of queries is $O(\frac{1}{\epsilon}n^2 \log n)$, which concludes the proof. □

Finally, there is another more involved randomized Local Search algorithm that produces a 0.4 approximation[46]. The main idea is to construct randomly the influence set A. However, the best we can do with a constant bias is to include with probability $1/2$ each element of V and results in a $1/4$ approximation. The reason behind this fact is that the geometry of the probability space we construct is too uniform and not nigh too a target underline geometry where the measure is spread with preference over subsets of V where $f(S) \geq r \cdot OPT$. Therefore, the way to go around this fact is to find a more appropriate probability space. Now, to that end we are confined to consider product spaces(decision about elements are independent) because we do not now all the dependencies between variables(exponential number of queries).

To succesfully construct such a probability space we must exploit the submodularity property of our function, which is more of a “local set” property. Hence, our goal will be to construct a set A according to which we will select all the elements in A with probability q and all elements in $V \setminus A$ with another probability $p < q$. This is called the *multinlinear extension* of a function $f(A)$[130]. So, the algorithm incrementally constructs a “good” set A by estimating the marginal contribution of each node of being included or exluded from A in a manner similar to the Local Search algorithm we presented above. We state here only the theorem without proof.

Theorem 5 (FMV[46]). *There is a Stochastic Local Search Algorithm that is a $(\frac{2}{5} - O(1))$-approximation algorithm for maximizing nonnegative submodular functions.*

3.4 Other Related Work

All studies in the Influence Maximization context treated influence as an indirect measure of the revenue extracted from the adoption of a product and the influence models studied did not incorporate pricing dynamics. Social Influence in the context of Revenue Maximization was first studied by Hartline et.al. [66]. Since then much work has been done mainly towards two directions; the study of posted price mechanisms, where price discrimination is not allowed, and game theoretic settings, where buyers act strategically according to their perceived utility for the product, which depends on other buyers's decisions.

Arthur et.al [8] considered a model where the seller does not have the freedom to approach potential buyers and only recommendations about products cascade through the network from an initial seed of early adopters. They showed that a very simple strategy achieves a constant approximation ratio.

Akghalpour et.al. [2] studied iterative posted price strategies where all interested buyers can by the product at the same price at a given time. They considered two different models allowing different re–pricing rates. They showed that when we are allowed to reprice very frequently the problem is inapproximable and provided approximation algorithms for a symmetric setting. On the other hand when re–pricing is allowed only at a limited rate they provided a FPTAS.

On a related work, Anari et.al. [5] considered a posted price setting where the product exhibits a special kind of externality, the so called historical externality. In their model buyers face the dilemma whether to buy a sub-par early version of the product or wait for a fully functional one. They assume that the seller commits a-priori to a pricing trajectory and then buyers decide when to buy the product. They considered that there are types buyers exhibiting different behaviour. In this setting they study existence and uniqueness of equilibria and provide an FPTAS to compute an approximate revenue maximizing pricing trajectory for two special cases.

Chen, Lu et.al [31] consider a case where each buyer's utility for the product depends on the set of buyers that own the product. They provide a game theoretic solution concept and consider Nash as well as Bayesian-Nash equilibria. In their setting where multiple equilibria might exist, they study two special cases, pessimistic and optimistic equilibria. They provide an algorithm to compute these equilibria and along the way find the price that maximized the sellers revenue.

Candogan, Bimpikis, Ozdaglar [25] considered a setting where a monopolist sells a divisible good to consumers that experience a positive network effect. They provide a game theoretic solution concept. They considered a two stage game where initially the seller sets an individual price for each consumer and then the buyers decide on the consumption, according to a quadratic utility function that depends on the consumption levels of their neighbours. They showed that when perfect price discrimination is allowed the optimal price depends linearly with the Bonanich Centrality of a buyer, whereas when only two prices are allowed the problem is reduced to Max-Cut.

Haghpanah et al.[65] study positive externalities in the context of designing Optimal Auctions. The consider single-parameter submodular network externalities in which a bidder's value for an outcome is a fixed private type times a known submodular function of the allocation of his friends. They prove that the optimal auction is APX-Hard, even on average. For a special kind of step-function externalities, they provide constant approximation algorithms. Moreover, using the influence and exploit idea, they use the algorithms developed by [46] to design constant approximation auction mechanisms.

Chapter 4

Preliminaries

Hartline et al.[66] considered a very general model on how externalities influence buyers' valuations of some product. However, studying the model in the most general(submodular) setting yields little insight in the combinatorial structure of the problem and the particularities of Revenue Maximization as opposed to other instances of submodular maximization. In this thesis, we focus on the Uniform Additive Model which provides the optimal trade-off between model complexity and analytical tractability. In this and the following chapters, we discuss the Revenue Maximization Problem under the Uniform Additive Model.

4.1 Uniform Additive Model

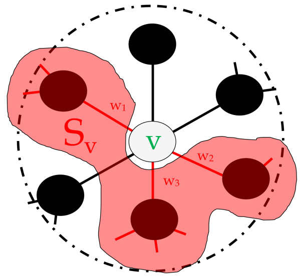
In our setting we are considering only positive externalities, that is the valuation of a buyers for the product can only increase if one of his social contacts already owns the product. The Uniform Additive Model is a way to model the extent that buyers influence each other in that respect. The model aims in capturing two aspects of the influence between buyers, dependence and uncertainty.

Towards that direction, in order to model each aspect certain assumptions are made. Buyers are influenced by each other pairwise. That is, the influence that a certain person exerts on someone is independent of the influence from another person. This assumption implies linearity of influence. Furthermore, since in reality buyers' valuations are private information and only rough estimations can be made about their true value, the model makes the simplistic assumption that valuations are uniformly distributed in an interval of possible values.

The Model We, thus, considered that we are given a (possibly directed) weighted social network $G(V, E, w)$ on the set V of potential buyers. For each edge $(i, j) \in E$, there is an associated positive weight w_{ij} and if $(i, j) \notin E$ we assume that $w_{ij} = 0$. A social network is undirected(or symmetric) if $w_{ij} = w_{ji}$ for all $i, j \in V$, and directed otherwise. Additionally, there may exist a non-negative weight w_{ii} associated with each buyer i[1]. These weights capture the pairwise influence between buyers.

[1]For directed networks we can safely ignore such weights, since we can consider an extra node $i^{'}$ for each such node with a single directed edge with weight $w_{i^{'}i} = w_{ii}$.

We consider that each buyer i has a valuation of the product $v_i(S)$ that depends only on the set S of buyers that already own the product. Particularly, we assume that $v_i(S)$ is a random variable distributed uniformly in the interval $[0, M_{i,S}]$, where $M_{i,S} = \sum_{j \in S} w_{ji}$ is the the total influence perceived by i from his neighbours that own the product.



Connection with LTM Alternatively, we can interpret the above process as each buyer picking a threshold θ_i uniformly at random from $[0, 1]$. Then if the offer x_i is smaller than $\theta_i M_{i,S}$, the buyer accepts, otherwise rejects the offer. The similarity with the Linear Threshold Model[76] now becomes evident. The only difference is that each buyer is considered only once at a chosen time, whilst in the LTM whenever the threshold is exceeded the node is activated. It would be interesting to generalize the LTM for a Posted Price Setting.

Myopic Pricing If we disregard the influence of a buyer i on future buyers, we can select a price x such that the expected revenue from buyer i is maximized. In the case of the UAM, this price is $M_{i,S}/2$ and buyer i has $1/2$ probability of accepting it. This pricing is not optimal because it disregards the effect that the buyer's decision may have on future buyers. Thus we call such a pricing strategy myopic.

4.2 Marketing Strategies and Revenue Maximization

In the setting we describe above, there is a seller wishing to exploit these externalities in order to devise a marketing strategy that will maximize his revenue. We consider that the seller has the freedom to approach each buyer individually and offer him a price. Therefore, a marketing strategy $(\pi, \mathbf{x})$ consists of a permutation of buyers V and a pricing vector $\mathbf{x} = (x_1, \ldots, x_n)$, where π determines the order by which buyers are visited and $\mathbf{x}$ the prices offered to them.

Note that the price that is offered to each buyer i depends on the set of buyers S that already own the product, which itself is a random variable. Consequently, we cannot calculate a priori the exact price x_i because it must depend on the S. We get around this fact by making the following observation. We observe that for any buyer i and any probability p that i accepts an offer, there is an (essentially unique) price x_p such that i accepts an offer of x_p with probability p. For the Uniform Additive Model, $x_p = (1-p)M_{i,S}$ and the expected revenue extracted from buyer i with such an offer is $p \cdot (1-p)M_{i,S}$.

Pricing with Probabilities Throughout this paper, we equivalently regard marketing strategies as consisting of a permutation π of the buyers and a vector $\mathbf{p} = (p_1, \ldots, p_n)$ of pricing probabilities. We note that if $p_i = 1$, i gets the product for free, while if $p_i = 1/2$, the price offered to i is (the myopic price of) $M_{i,S}/2$. We assume that $p_i \in [1/2, 1]$, since any expected revenue in $[0, M_{i,S}/4]$ can be achieved with such pricing probabilities. The expected revenue of a marketing strategy $(\pi, \mathbf{p})$ is:

$$R(\pi, \mathbf{p}) = \sum_{i \in V} p_i(1-p_i) \left(\sum_{\pi(j) < \pi(i)} w_{j,i} p_j + w_{ii} \right) \tag{4.1}$$

The problem of Revenue Maximization under the Uniform Additive Model is to find a marketing strategy $(\pi, \mathbf{p})$ that extracts a maximum revenue of $R(\pi^*, \mathbf{p}^*)$ from a given social network $G(V, E, w)$.

Bounds on the Maximum Revenue Let $N = \sum_{i \in V} w_{ii}$ and $W = \sum_{i<j} w_{ij}$, if the social network G is undirected, and $W = \sum_{(i,j) \in E} w_{ij}$, if G is directed. Then an upper bound on the maximum revenue of G is $R^* = (W + N)/4$, and follows by summing up the myopic revenue over all edges of G [66, Fact 1]. For a lower bound on the maximum revenue, if G is undirected (resp. directed), approaching the buyers in any order (resp. in a random order) and offering them the myopic price yields a revenue of $(W + 2N)/8$ (resp. $(W + 4N)/16$). Thus, myopic pricing achieves an approximation ratio of 0.5 for undirected networks and of 0.25 for directed networks. nd of 0.25 for directed networks.

4.3 Ordering and NP-Hardness

Revenue maximization exhibits a dual nature involving optimizing both the pricing probabilities and the sequence of offers. For directed networks, finding a good ordering $\vec{\pi}$ of the buyers bears a resemblance to the Maximum Acyclic Subgraph problem, where given a directed network $G(V, E, w)$, we seek for an acyclic subgraph of maximum total edge weight. In fact, any permutation $\vec{\pi}$ of V corresponds to an acyclic subgraph of G that includes all edges going forward in $\vec{\pi}$, i.e, all edges (i, j) with $\pi_i < \pi_j$. [66, Lemma 3.2] shows that given a directed network G and a pricing probability vector $\vec{p}$, computing an optimal ordering of the buyers (for the particular $\vec{p}$) is equivalent to computing a Maximum Acyclic Subgraph of G, with each edge (i, j) having a weight of $p_i p_j (1 - p_j) w_{ij}$. Consequently, computing an ordering $\vec{\pi}$ that maximizes $R(\vec{\pi}, \vec{p})$ is **NP**-hard and Unique-Games-hard to approximate within a factor greater than 0.5 [64].

On the other hand, we show that in the undirected case, if the pricing probabilities are given, we can easily compute the best ordering of the buyers.

Lemma 4. *Let $G(V, E, w)$ be an undirected social network, and let $\vec{p}$ be any pricing probability vector. Then, approaching the buyers in non-increasing order of their pricing probabilities maximizes the revenue extracted from G under $\vec{p}$.*

Proof. We consider an optimal ordering $\vec{\pi}$ (wrt. $\vec{p}$) that minimizes the number of buyers' pairs appearing in increasing order of their pricing probabilities, namely, the number of pairs i_1, i_2 with $p_{i_1} < p_{i_2}$ and $\pi_{i_1} < \pi_{i_2}$. If there is such a pair in $\vec{\pi}$, we can find a pair of buyers i and j with $p_i < p_j$ such that i appears just before j in $\vec{\pi}$. Then, switching the positions of i and j in $\vec{\pi}$ changes the expected revenue extracted from G under $\vec{p}$ by $p_i p_j w_{ij}(p_j - p_i) \geq 0$, a contradiction. □

A consequence of Lemma 4 is that [66, Lemma 3.2] does not imply the **NP**-hardness of revenue maximization for undirected social networks. Hence, there are two natural questions that we may ask. Does the ordering play any role in extracting revenue from the network or given an ordering we can always find a pricing vector $\mathbf{p}$ that manages to extract the maximal amount of revenue?

To answer this question, we provide an example where fixing the ordering dwindles our ability to extract the optimal revenue. e consider an (undirected) simple cycle with 4 nodes, numbered as they appear on the cycle, and unit weights on its edges. Proposition 5 shows

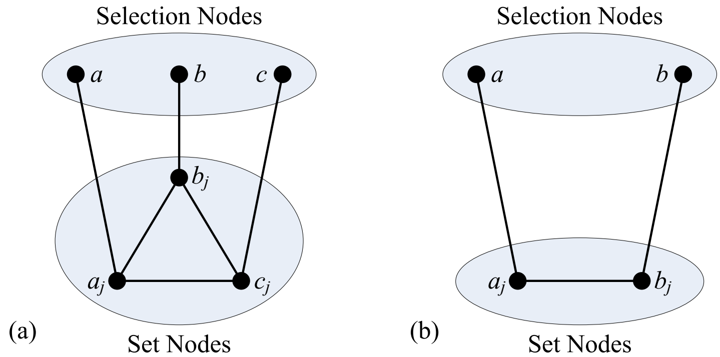


Figure 4.1: Examples of (a) an extended triangle and (b) a 3-path, used in the proof of Lemma 5. We create an extended triangle for each 3-item set T_j and a 3-path for each 2-item set T_j. The set nodes are different for each set T_j, while the selection nodes are common for all sets.

that the optimal ordering is $(1, 3, 2, 4)$, the optimal pricing vector is $(1, 0.5, 1, 0.5)$, and the maximum revenue is 1. On the other hand, if the nodes are ordered as they appear in the cycle, i.e., as in $(1, 2, 3, 4)$, the optimal pricing vector is $(1, \sqrt{2}/2, (1 + \sqrt{2})/2, 0.5)$, and the resulting revenue is 0.7772.

So, given that the ordering does have an impact after all, what can we say about the hardness of the problem in the undirected setting. The following lemma employs a reduction from monotone One-in-Three 3-SAT, and shows that revenue maximization is **NP**-hard for undirected networks.

Lemma 5. *The problem of computing a marketing strategy that extracts the maximum revenue from an undirected social network is* **NP***-hard.*

Proof. In monotone One-in-Three 3-SAT, we are given a set V of n items and m subsets $T_1, \dots, T_m$ of V, with $2 \leq |T_j| \leq 3$ for each $j \in \{1, \dots, m\}$. We ask for a subset $S \subset V$ such that $|S \cap T_j| = 1$ for all $j \in \{1, \dots, m\}$. Monotone One-in-Three 3-SAT is shown **NP**-complete in [**?**]. In the following, we let m_2 (resp. m_3) denote the number of 2-item (resp. 3-item) sets T_j in an instance $(V, T_1, \dots, T_m)$ of monotone One-in-Three 3-SAT.

Given $(V, T_1, \dots, T_m)$, we construct an undirected social network G. The network G contains a *selection-node* corresponding to each item in V. There are no edges between the selection nodes of G. For each 3-item set $T_j = \{a, b, c\}$, we create an *extended triangle* consisting of a triangle on three *set nodes* a_j, b_j, and c_j, and three additional edges that connect a_j, b_j, c_j to the corresponding selection nodes a, b, and c (see also Fig. 4.1.a). For each 2-item set $T_j = \{a, b\}$, we create a 3-*path* consisting of an edge connecting two set nodes a_j and b_j, and two additional edges connecting a_j and b_j to the corresponding selection nodes a and b (see also Fig. 4.1.b). Therefore, G contains $n + 2m_2 + 3m_3$ nodes and $3m_2 + 6m_3$ edges. The weight of all edges of G is 1. We next show that $(V, T_1, \dots, T_m)$ is a YES-instance of monotone One-in-Three 3-SAT iff the maximum revenue of G is at least $\frac{177}{128}\, m_3 + \frac{3}{4}\, m_2$.

By Lemma 4, the revenue extracted from G is maximized if the nodes are approached in non-increasing order of their pricing probabilities. Therefore, we can ignore the ordering of the nodes, and focus on their pricing probabilities. The important property is that if each extended triangle (Fig. 4.1.a) is considered alone, its maximum revenue is 177/128, and is obtained when exactly one of the selection nodes a, b, c has a pricing probability of 1/2 and the other two have a pricing probability of 1. More specifically, since the selection nodes a, b, c have degree 1, the revenue of the extended triangle is maximized when they have a pricing probability of either 1 or 1/2. If all a, b, c have a pricing probability of 1, the best revenue of the extended triangle is ≈ 1.196435, and is obtained when one of a_j, b_j, and c_j has a pricing probability of ≈ 0.7474, the other has a pricing probability of ≈ 0.5715, and the third has a pricing probability of 1/2. If all a, b, c have a pricing probability of 1/2, the best revenue of the extended triangle is again ≈ 1.196435, and is obtained with the same pricing probabilities of a_j, b_j, and c_j. If two of a, b, c (say a and b) have a pricing probability of 1/2 and c has a pricing probability of 1, the best revenue of the extended triangle is $\frac{21}{16} = 1.3125$, and is obtained when one of a_j and b_j has a pricing probability of 1, the other has a pricing probability of 3/4, and c_j has a pricing probability of 1/2. Finally, if two of a, b, c (say b and c) have a pricing probability of 1 and a has a pricing probability of 1/2, we extract a maximum revenue from the extended triangle, which is $\frac{177}{128} = 1.3828125$ and is obtained when a_j has a pricing probability of 1, one of b_j and c_j has a pricing probability of 9/16, and the other has a pricing probability of 1/2.

Similarly, if each 3-path (Fig. 4.1.b) is considered alone, its maximum revenue is 3/4, and is obtained when exactly one of the selection nodes a, b has a pricing probability of 1/2 and the other has a pricing probability of 1. In fact, since the 3-path is a bipartite graph, Proposition 5 implies that the maximum revenue, which is 3/4, is extracted when a_j and b have a pricing probability of 1 and b_j and a have a pricing probability of 1/2 (or the other way around). If both a and b have a pricing probability of 1, the best revenue of the 3-path is 41/64 and is obtained when one of a_j and b_j has a pricing probability of 5/8, and the other has a pricing probability of 1/2. If both a and b have a pricing probability of 1/2, the best revenue of 3-path is again 41/64 and is obtained when one of a_j and b_j has a pricing probability of 1, and the other has a pricing probability of 5/8.

If $(V, T_1, \ldots, T_m)$ is a YES-instance of monotone One-in-Three 3-SAT, we assign a pricing probability of 1/2 to the selection nodes in S and a pricing probability of 1 to the selection nodes in $V \setminus S$, where S is a set with exactly one element of each T_j. Thus, we have exactly one selection node with pricing probability 1/2 in each extended triangle and in each 3-path. Then, we can set the pricing probabilities of the set nodes as above, so that the revenue of each extended triangle is 177/128 and the revenue of each 3-path is 3/4. Thus, the maximum revenue of G is at least $\frac{177}{128} m_3 + \frac{3}{4} m_2$.

For the converse, we recall that the edges of G can be partitioned into m_3 extended triangles and m_2 3-paths. Consequently, if the maximum revenue of G is at least $\frac{177}{128} m_3 + \frac{3}{4} m_2$, each extended triangle contributes exactly 177/128 and each 3-path contributes exactly 3/4 to the revenue of G. Thus, by the analysis on their revenue above, each extended triangle and each 3-path includes exactly one selection node with a pricing probability of 1/2. Therefore, if we let S consist of the selection nodes with pricing probability 1/2, we have that $|S \cap T_j| = 1$ for all $j \in \{1, \ldots, m\}$. □

The fact that the problem remains **NP**-hard in the undirected setting suggests that the pricing aspect lies at the heart of the problem. The very proof of **NP**-hardness illustrates that selecting prices for vertices among the continuous space $[1/2, 1]^n$ in an optimal way encodes complex information. In the sense that assigning prices to vertices essentially corresponds to finding an optimal weighted multi-cut(in our proof the resemblance with Max-cut is more apparent). That is, we could consider a discretization of the pricing range $[1/2, 1]$ in sufficiently many k prices and ask the question, how we should distribute vertices in these pricing classes such that corresponding weighted multi-cut is maximized? In the next chapter we will exploit this idea to provide an improved approximation algorithm for the problem.

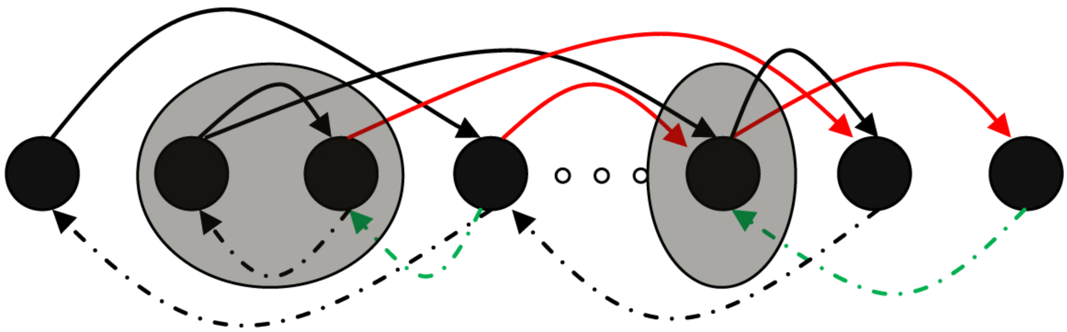


Figure 4.2: Illustration of the argument used in the proof of Lemma 1.

4.4 Determining the Sequence

In this section we aim to provide further connections between revenue maximization and the notion of multi-cuts in the network. Particularly, we will show how multi-cuts are related with the Maximum Feedback Arc set problem.

Definition 4. *The problem of Maximum Feedback Arc set is given a directed graph $G(V, E)$ find an ordering(permutation) of vertices such that the total weight of edges that are going from indices higher in the permutation to lower ones is maximal.*

This problem is equivalent to the maximum acyclic sub-graph problem and it is NP-Hard. However, as discussed before there exists a simple randomized 1/2-approximate algorithm, which just takes a random ordering. In 2008 Guruswami et.al [64] showed that the problem cannot be approximated better than 1/2 unless the Unique Games Conjecture [78, **?**] is false. These two facts imply that not only a random ordering is good but it is probably the best we can do up to $o(1)$ factors [27].

We proceed now with observations that will motivate later developments. Notice that the structure of Max-FAS has the feature that an optimal solution is optimal for the same problem when constrained to any subset of vertices that are consecutive in the optimal permutation. This means that if we knew that the only candidates for positions $x_k, \ldots, x_{k+m}$ are those in a set $|X| = m$, then we would solve this problem optimally without worrying for external interference. At this point we state the following structural lemma:

Lemma 6. *An optimal ordering of vertices for the MAX-FAS, "contains" the MAX-CUT solution. Meaning that if the optimal ordering is $\pi^* = (i_1, \ldots, i_n)$, then there is $k \leq n$ such that $S = \{i_k, \ldots, i_n\}$ and $T = V \setminus S = \{i_1, \ldots, i_{k-1}\}$ is an optimal solution for the MAX-CUT problem with vertices in V.*

Proof. We will prove it by contradiction. Assume we have an optimal solution and suppose that there isn't such k, then that would mean that there is a set S' with elements that are not all consecutive such that the directed cut between S' and $V \setminus S' = T'$ is an optimal solution for the MAX-CUT problem. If that is the case then we could keep the ordering inside the two sets S', T' and relocate S' in the end of the sequence. Now in terms of the weight of the FAS we created, we would miss the edges going from T' to S', but we would gain the edges from S' to T'. However, because S'andT' is an optimal solution for the MAX-CUT problem we are guaranteed to gain feedback arcs. But we assumed that the initial ordering was optimal. □

The algorithmic interpretation of the previous lemma is that if there is a unique maximum cut we can extract a partial ordering of vertices. Then due to the optimality of sub-solutions we could proceed and derive the total ordering in a divide and conquer manner.

Corollary 1. *Assume that there is a unique MAX-CUT solution for every subset of vertices in V, then an algorithm that solved optimally the directed MAX-CUT problem would also solve optimally the MAX-FAS problem.*

Let us look again closer to the 1/2 approximation algorithm for MAX-FAS. This algorithm just outputs a random permutation, however we can view this process of obtaining a random permutation as a synthesis of processes where the core process is the 1/2 approximate algorithm for the max-cut problem. Specifically, we can think the final permutation as a result of more refined partial orderings of vertices. Particularly, at each time point we would have disjoint sets of vertices $S_1, S_2, \ldots, S_k$ and a total ordering on these sets $S_1 \succ S_2 \succ \ldots \succ S_k$. So until we had only singleton sets we would further refine our ordering, by applying the simple algorithm for max-cut.

The random permutation algorithm can be seen as naively trying to approximate the problem by substituting an algorithm that solves optimally MAX-CUT, without even caring about the possible existence of many possible cuts and how to choose between them. Surprisingly enough, this algorithm is currently the state of the art. The question that arises is: "Instead of considering as the core process the simple approximation algorithm, wouldn't we get much better results if we applied the much more sophisticated approximation algorithm of Feige and Goemanns [45] for the directed MAX-CUT?". Obtaining a better algorithm for the MAX-FAS, though very unlikely, would have great implications as it would disprove the Unique Games Conjecture.

Chapter 5

Influence and Exploit

Influence and Exploit strategies are simple, elegant and manage to extract a significant fraction of the optimal revenue. However, their algorithmic properties are hardly well understood, namely their limits in approximating the optimal revenue and the computational hardness of finding the optimal revenue. In this chapter we manage to significantly extend the understanding of IE strategies.

Before, proceeding with the technical part, we briefly discuss the motivation behind considering such strategies. In the previous chapter, we saw that the uncertainty in the model resulted in a continuous search space were marketing strategies lie and the in tandem hardness. The Influence and Exploit idea just select two extreme points in the range and assigns vertices in pricing classes. The motivation behind this approach partly stems from the fact that the combinatorial properties of partitioning vertices into two sets have been extensively studied and partly to translate results from the theory of maximizing submodular set functions inspired from the Influence Maximization paradigm. Lastly, IE strategies are better suited for practical real world implementations, since many marketing efforts are usually of the discount/full price form and buyers are accustomed to it.

An *Influence-and-Exploit* (IE) strategy $\mathrm{IE}(A, p)$ consists of a set of buyers A receiving the product for free and a pricing probability p offered to the remaining buyers in $V \setminus A$, who are approached in a random order. We slightly abuse the notation and let $\mathrm{IE}(q, p)$ denote an IE strategy where each buyer is selected in A independently with probability q. $\mathrm{IE}(A, p)$ extracts an expected (wrt the random ordering of the exploit set) revenue of:

$$R_{\mathrm{IE}}(A, p) = p(1-p) \sum_{i \in V \setminus A} \left(w_{ii} + \sum_{j \in A} w_{ji} + \sum_{j \in V \setminus A,\ j \neq i} \frac{p\, w_{ji}}{2} \right) \tag{5.1}$$

Specifically, $\mathrm{IE}(A, p)$ extracts a revenue of $p(1-p)w_{ji}$ from each edge (j, i) with buyer j in the influence set A and buyer i in the exploit set $V \setminus A$. Moreover, $\mathrm{IE}(A, p)$ extracts a revenue of $p^2(1-p)w_{ji}$ from each edge (j, i) with both j, i in the exploit set, if j appears before i in the random order of $V \setminus A$, which happens with probability $1/2$.

The problem of finding the best IE strategy is to compute a subset of buyers A^* and a pricing probability p^* that extract a maximum revenue of $R_{\mathrm{IE}}(A^*, p^*)$ from a given social network $G(V, E, w)$.

Interestingly, even very simple IE strategies extract a significant fraction of the maximum revenue. For example, for undirected social networks, $R_{\mathrm{IE}}(\emptyset, 2/3) = (4W + 6N)/27$, and thus $\mathrm{IE}(\emptyset, 2/3)$ achieves an approximation ratio of $\frac{16}{27} \approx 0.592$. For directed networks, $R_{\mathrm{IE}}(\emptyset, 2/3) =$

$(2W + 6N)/27$, and thus IE$(\emptyset, 2/3)$ achieves an approximation ratio of $\frac{8}{27} \approx 0.296$.

5.1 Optimal Influence and Exploit Strategies

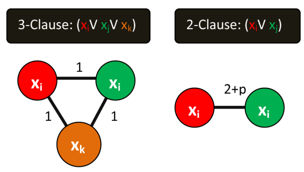
Influence and Exploit is a powerful idea and yields good algorithms approximating the optimal revenue. Does the restriction of vertices only in two pricing classes makes the problem easier from a computational complexity perspective? Intuitively, this should not be true as we saw that hardness in the undirected setting was established by using a generalized notion of max-cut, which is very close to the Influence and Exploit idea. Then, if the restricted problem is as hard as the original, what can be said about the potential of IE strategies in approximating the optimal revenue? Meaning, can the optimal revenue be achieved for the general case by using only two prices? Such, a result would be surprising as it would imply a surprising degeneracy in the model.

In this section we provide results about the computational hardness of finding the optimal IE strategy. Furthermore, we provide lower bounds on the performance of the Optimal IE strategy in terms of approximating the optimal revenue as well as corresponding upper bounds for the directed case.

Hardness The following lemma employs a reduction from monotone One-in-Three 3-SAT[122], and shows that computing the best IE strategy is **NP**-hard.

Lemma 7. *Let $p \in [1/2, 1)$ be any fixed pricing probability. The problem of finding the best IE strategy with pricing probability p is* **NP***-hard, even for undirected social networks.*

Proof. We recall that in monotone One-in-Three 3-SAT, we are given a set V of n items and m subsets $T_1, \dots, T_m$ of V, with $2 \le |T_j| \le 3$ for each $j \in \{1, \dots, m\}$. We ask for a subset $S \subset V$ such that $|S \cap T_j| = 1$ for all $j \in \{1, \dots, m\}$. Given $(V, T_1, \dots, T_m)$, we construct an undirected social network G on V.



For each 3-item set $T_j = \{a, b, c\}$, we create a *set-triangle* on nodes a, b, and c with 3 edges of weight 1. For each 2-item set $T_j = \{a, b\}$, we add a *set-edge* $\{a, b\}$ of weight $2 + p$, where p is the pricing probability. To avoid multiple appearances of the same edge, we let the weight of each edge be the total weight of its appearances. Namely, if an edge e appears in k_3 set-triangles and in k_2 set-edges, e's weight is $k_3 + (2 + p)k_2$. We observe that for any $p \in [1/2, 1)$, the maximum revenue extracted from any set-triangle and any set-edge is $p(1-p)(2+p)$, by giving the product for free to exactly one of the nodes of the set-triangle (resp. the set-edge).

We next show that $(V, T_1, \dots, T_m)$ is a YES-instance of monotone One-in-Three 3-SAT iff there is an influence set A in G such that $R_{\text{IE}}(A, p) \ge mp(1-p)(2+p)$. If $(V, T_1, \dots, T_m)$ is a YES-instance of monotone One-in-Three 3-SAT, we let the influence set $A = S$, where S is a set with exactly one element of each T_j. Then, we extract an expected revenue of $p(1-p)(2+p)$ from each set-triangle and each set-edge in G, which yields an expected revenue of $mp(1-p)(2+p)$ in total. For the converse, if there is an influence set A in G such that $R_{\text{IE}}(A, p) \ge mp(1-p)(2+p)$, we let $S = A$. Since $R_{\text{IE}}(A, p) \ge mp(1-p)(2+p)$, and since the edges of G can be partitioned into m set-triangles and set-edges, each with a maximum revenue of at most $p(1-p)(2+p)$, each set-triangle and each set-edge contributes exactly $p(1-p)(2+p)$ to $R_{\text{IE}}(A, p)$. Therefore, for all set-triangles and all set-edges, there is exactly one node in A. Thus, we have that $|S \cap T_j| = 1$ for all $j \in \{1, \dots, m\}$. ☐

Efficiency of Influence and Exploit Strategies Given the hardness of computing both the optimal marketing strategy and the optimal IE strategy, how can we compare the two revenues? The usual direction would be to indirectly compare them by obtaining an upper bound for the optimal strategy and a lower bound for the performance of the optimal IE strategy. However, we only have the naive upper bounds discussed earlier which can be arbitrarily away from the optimal revenue.

To come around this obstacle we exploit the fact that the revenue of any strategy in the Uniform Additive Model can be written as a linear function of the edge weights, where each weight is multiplied by a factor depending on the probability that the two respective vertices accept the offer, thus utilizing Local Analysis. A lower bound on the approximation ratio can be derived by comparing termwise the two revenues and keeping the worst ratio accross different terms(w.r.t the edges). Hence, we show that the best IE strategy manages to extract a significant fraction of the maximum revenue.

Theorem 6. *For any undirected social network, there is an IE strategy with pricing probability* 0.586 *whose revenue is at least* 0.9111 *times the maximum revenue.*

Proof. We consider an arbitrary undirected social network $G(V, E, w)$. We assume that the optimal IE strategy has at its disposal the optimal probability vector $\mathbf{p}^*$(corresponding to the optimal marketing strategy) and assigns vertices in to one of the two classes accordingly. A lower bound on the optimal IE $\text{IE}(A, \hat{p})$ strategy can be given by any specific decision rule. We consider that this decision is madeby applying randomized rounding to $\mathbf{p}^*$. We show that for $\hat{p} = 0.586$, the expected (wrt the randomized rounding choices) revenue of $\text{IE}(A, \hat{p})$ is at least 0.9111 times the revenue extracted from G by the best ordering for $\mathbf{p}$ (recall that by Lemma 4, the best ordering is to approach the buyers in non-increasing order of their pricing probabilities).

Without loss of generality, we assume that $p_1 \geq p_2 \geq \cdots \geq p_n$, and let π be the identity permutation. Then, $R(\pi, \mathbf{p}) = \sum_{i \in V} p_i(1 - p_i)w_{ii} + \sum_{i<j} p_i p_j (1 - p_j)w_{ij}$.

For the IE strategy, we assign each buyer i to the influence set A independently with probability $I(p_i) = \alpha(p_i - 0.5)$, for some appropriate $\alpha \in [0, 2]$, and to the exploit set with probability $E(p_i) = 1 - I(p_i)$. By linearity of expectation, the expected revenue of $\text{IE}(A, \hat{p})$ is:

$$
\begin{aligned}
R_{\text{IE}}(A, \hat{p}) &= \sum_{i<j} \hat{p}(1 - \hat{p})(I(p_i)E(p_j) + E(p_i)I(p_j) + \hat{p}\, E(p_i)E(p_j))w_{ij} \\
&+ \sum_{i \in V} \hat{p}(1 - \hat{p})E(p_i)w_{ii}
\end{aligned}
$$

Specifically, $\text{IE}(A, \hat{p})$ extracts a revenue of $\hat{p}(1 - \hat{p})w_{ii}$ from edge loop $\{i, i\}$, if i is included in the exploit set. Moreover, $\text{IE}(A, \hat{p})$ extracts a revenue of $\hat{p}(1 - \hat{p})w_{ij}$ from each edge $\{i, j\}$, $i < j$, if one of i, j is included in the influence set A and the other is not, and a revenue of $\hat{p}^2(1 - \hat{p})w_{ij}$ if both i and j are included in the exploit set $V \setminus A$ (note that the order in which i and j are considered is insignificant).

The approximation ratio is derived as the minimum ratio between any pair of terms in $R(\pi, \mathbf{p})$ and $R_{\text{IE}}(A, \hat{p})$ corresponding to the same loop $\{i, i\}$ or to the same edge $\{i, j\}$. For a weaker bound, we observe that for $\alpha = 1.43$ and $\hat{p} = 0.586$, both

$$
\begin{aligned}
&\min\nolimits_{0.5 \leq x \leq 1} \quad \frac{\hat{p}\,(1 - \hat{p})\,E(x)}{x\,(1 - x)} \quad \text{and} \\
&\min\nolimits_{0.5 \leq y \leq x \leq 1} \quad \frac{\hat{p}\,(1 - \hat{p})(I(x)\,E(y) + E(x)\,I(y) + \hat{p}\,E(x)\,E(y))}{x\,y\,(1 - y)}
\end{aligned}
\tag{5.2}
$$

are at least 0.8024. More precisely, the former quantity is minimized for $x \approx 0.7104$, for which it becomes ≈ 0.8244. For any fixed value of $y \in [0.5, 1.0]$, the latter quantity is minimized for $x = 1.0$. The minimum value is 0.8024 for $x = 1.0$ and $y \approx 0.629$.

For the stronger bound of 0.9111, we let $\hat{p} = 0.586$, and for each buyer i, let the rounding parameter $\alpha(p_i)$ be chosen according to the following piecewise linear function of p_i :

$$\alpha(p_i) = \begin{cases} 5.0\,(p_i - 0.5) & \text{if } 0.5 \le p_i \le 0.7 \\ 1.0 + 3.3\,(p_i - 0.7) & \text{if } 0.7 < p_i \le 0.8 \\ 1.33 + 3.0\,(p_i - 0.8) & \text{if } 0.8 < p_i \le 0.9 \\ 1.63 + 3.7\,(p_i - 0.9) & \text{if } 0.9 < p_i \le 1.0 \end{cases}$$

The quantity on the left of (5.2) is minimized for $x = 0.8$, for which it becomes ≈ 0.9112. For any fixed $x \in [0.5, 0.949]$, the quantity on the right of (5.2) is minimized for $y = 0.5$. The minimum value is 0.9111 for $x \approx 0.7924$ and $y = 0.5$. For any $x \in (0.949, 0.983]$, the latter quantity is minimized for $y = 0.7$. The minimum value, over all $x \in (0.949, 0.983]$, is ≈ 0.93 at $x = 0.983$ and $y = 0.7$. For any fixed $x \in (0.983, 1.0]$, the quantity on the right of (5.2) is minimized for some $y \in [0.7, 0.8]$. Moreover, for all $y \in [0.7, 0.8]$, this quantity is minimized for $x = 1.0$. The minimum value is ≈ 0.9112 at $x = 1.0$ and $y \approx 0.8$. □

Theorem 7. *For any directed social network, there is an IE strategy with pricing probability* $2/3$ *whose expected revenue is at least* 0.55289 *times the maximum revenue.*

Proof. As before, we consider an arbitrary directed social network $G(V, E, w)$, start from an arbitrary pricing probability vector $\mathbf{p}$, and obtain an IE strategy $\mathrm{IE}(A, \hat{p})$ by applying randomized rounding to $\mathbf{p}$. We show that for $\hat{p} = 2/3$, the expected (wrt the randomized rounding choices) revenue of $\mathrm{IE}(A, \hat{p})$ is at least 0.55289 times the revenue extracted from G under the best ordering for $\mathbf{p}$ (which ordering is Unique-Games-hard to approximate within a factor less than 0.5!).

We recall that in the directed case, we can, without loss of generality, ignore loops (i, i). Let π be the best ordering π for $\mathbf{p}$. Then, the maximum revenue extracted from G with pricing probabilities $\mathbf{p}$ is $R(\pi, \mathbf{p}) \le \sum_{(i,j)\in E} p_i p_j (1 - p_j) w_{ij}$.

As in the proof of Theorem 6, we assign each buyer i to the influence set A independently with probability $I(p_i) = \alpha(p_i - 0.5)$, for some $\alpha \in [0, 2]$, and to the exploit set with probability $E(p_i) = 1 - I(p_i)$. By linearity of expectation, the expected (wrt the randomized rounding choices) revenue extracted by $\mathrm{IE}(A, \hat{p})$ is:

$$R_{\mathrm{IE}}(A, \hat{p}) = \sum_{(i,j)\in E} \hat{p}(1 - \hat{p})(I(p_i)E(p_j) + 0.5\,\hat{p}\,E(p_i)E(p_j))w_{ij}$$

Specifically, $\mathrm{IE}(A, \hat{p})$ extracts a revenue of $\hat{p}(1 - \hat{p})w_{ij}$ from each edge (i, j), if i is included in the influence set and j is included in the exploit set, and a revenue of $\hat{p}^2(1 - \hat{p})w_{ij}$ if both i and j are included in the exploit set $V \setminus A$ and i appears before j in the random order of $V \setminus A$.

The approximation ratio is derived as the minimum ratio between any pair of terms in $R(\pi, \mathbf{p})$ and $R_{\mathrm{IE}}(A, \hat{p})$ corresponding to the same edge (i, j). Thus, we select $\hat{p}$ and α so that the following quantity is maximized:

$$\min_{0.5 \le x, y \le 1} \frac{\hat{p}\,(1 - \hat{p})(I(x)\,E(y) + 0.5\,\hat{p}\,E(x)\,E(y))}{x\,y\,(1 - y)}$$

We observe that for $\hat{p} = 2/3$ and $\alpha = 1.0$, this quantity is simplified to $\min_{y \in [0.5,1]} \frac{2(3-2y)}{27y(1-y)}$. The minimum value is ≈ 0.55289 at $y = \frac{3-\sqrt{3}}{2}$. □

Similarly, we can show that there is an IE strategy with pricing probability 1/2 whose revenue is at least 0.8857 (resp. 0.4594) times the maximum revenue for undirected (resp. directed) networks.

Approximability of the Maximum Revenue for Directed Networks The results of [66, Lemma 3.2] and [64] suggest that given a pricing probability vector $\mathbf{p}$, it is Unique-Games-hard to compute a vertex ordering π of a directed network G for which the revenue of $(\pi, \mathbf{p})$ is at least 0.5 times the maximum revenue of G under $\mathbf{p}$. An interesting consequence of Theorem 7 is that this inapproximability bound of 0.5 does not apply to revenue maximization in the Uniform Additive Model. In particular, given a pricing probability vector $\mathbf{p}$, Theorem 7 constructs, in linear time, an IE strategy with an expected revenue of at least 0.55289 times the maximum revenue of G under $\mathbf{p}$. This does not contradict the results of [?, 64], because the pricing probabilities of the IE strategy are different from $\mathbf{p}$. Moreover, in the Uniform Additive Model, different acyclic (sub)graphs (equivalently, different vertex orderings) allow for a different fraction of their edge weight to be translated into revenue (for an example, see Section ??, in the Appendix), while in the reduction of [66, Lemma 3.2], the weight of each edge in an acyclic subgraph is equal to its revenue.

We show a simple example where different acyclic subgraphs (equivalently, different vertex orderings) of the social network allow for a different fraction of their edge weight to be translated into revenue. To this end, we consider a simple directed network G on $V = \{u_1, u_2, u_3, u_4\}$. G contains an edge from each vertex u_i to each vertex u_j with $j > i$, that is 6 edges in total. Formally, $E = \{(u_i, u_j) : 1 \leq i < j \leq 4\}$. The weight of each edge is 1.

In ordering $\pi_1 = (u_1, u_2, u_3, u_4)$, all edges go forward. So, π_1 corresponds to an acyclic subgraph with edge weight 6. The optimal pricing probabilities for π_1 are $\mathbf{p}_1 = (1, 0.7474, 0.5715, 0.5)$ and extract a revenue of $R(\pi_1, \mathbf{p}_1) = 1.1964$ from G. Thus, π_1 allows for a revenue equal to 19.943% of its edge weight.

Similarly, ordering $\pi_2 = (u_1, u_3, u_2, u_4)$ corresponds to an acyclic subgraph with edge weight 5. The optimal pricing probabilities for π_2 are $\mathbf{p}_2 = (1, 0.625, 0.625, 0.5)$ and extract a revenue of $R(\pi_2, \mathbf{p}_2) = 1.03125$. So, π_2 allows for a revenue equal to 20.625% of its edge weight.

Ordering $\pi_3 = (u_2, u_1, u_3, u_4)$ also corresponds to an acyclic subgraph with edge weight 5. The optimal pricing probabilities for π_3 are $\mathbf{p}_3 = (1, 1, 0.5625, 0.5)$ and extract a revenue of $R(\pi_3, \mathbf{p}_3) = 1.1328$. Thus, π_3 allows for revenue equal to 22.656% of its edge weight. Also, the revenue extracted by IE$(\{u_1, u_2\}, 0.5147)$ is 1.0634. Thus, π_3 allows for an IE strategy extracting a revenue equal to 21.268% of its edge weight.

IE$(\{u_1, u_2\}, 0.5147)$, for example, approximates the maximum revenue of G within a factor of $\frac{1.0634}{1.1964} \approx 0.8888$. On the other hand, if we consider a random ordering of u_1 and u_2 and of u_3 and u_4, we obtain a vertex ordering π', which combined with $\mathbf{p}_1$, gives an expected revenue of ≈ 1.0306. Hence, $(\pi', \mathbf{p})$ approximates the maximum revenue of G under $\mathbf{p}_1$ within a factor of $\frac{1.0306}{1.1964} \approx 0.8614$. On the other hand, π' defines an acyclic subgraph of G which has an expected edge weight of 5 and approximates the edge weight of the maximum acyclic subgraph of G within a factor of $\frac{5}{6} \approx 0.8333$.

Thus, although the IE strategy of Theorem 7 is 0.55289-approximate with respect to the maximum revenue of G under $\mathbf{p}$, its vertex ordering combined with $\mathbf{p}$ may generate a revenue of less than 0.5 times the maximum revenue of G under $\mathbf{p}$. In fact, based on Theorem 7, we obtain, in Section **??**, a polynomial-time algorithm that approximates the maximum revenue of a directed network G within a factor of 0.5011.

The following propositions establish a pair of inapproximabity results for revenue maximization in the Uniform Additive Model.

Proposition 1. *Assuming the Unique Games conjecture, it is* **NP***-hard to compute an IE strategy with pricing probability* $2/3$ *that approximates within a factor greater than* $3/4$ *the maximum revenue of a directed social network in the Uniform Additive Model.*

Proof. Let $G(V, E, w)$ be a directed social network, and let π^* be a vertex ordering corresponding to an acyclic subgraph of G with a maximum edge weight of W^*. Then, approaching the buyers according to π^* and offering a pricing probability of $2/3$ to each of them, we extract a revenue of $4W^*/27$. Therefore, the maximum revenue of G is at least $4W^*/27$.

Now, we assume an influence set A so that $\mathrm{IE}(A, 2/3)$ approximates the maximum revenue of G within a factor of r. Thus, $R_{\mathrm{IE}}(A, 2/3) \geq 4rW^*/27$. Let π be the order in which $\mathrm{IE}(A, 2/3)$ approaches the buyers, and let (i, j) be any edge with $\pi_i < \pi_j$, namely, any edge from which $\mathrm{IE}(A, 2/3)$ extracts some revenue. Since the revenue extracted from each such edge (i, j) is at most $2w_{ij}/9$, the edge weight of the acyclic subgraph defined by π is at least $\frac{9}{2}R_{\mathrm{IE}}(A, 2/3) \geq \frac{2r}{3}W^*$.

Hence, given an r-approximate $\mathrm{IE}(A, 2/3)$, we can approximate W^* within a ratio of $2r/3$. The proposition follows from [64, Theorem 1.1], which assumes the Unique Games conjecture and shows that it is **NP**-hard to approximate W^* within a ratio greater than $1/2$. □

Proposition 2. *Assuming the Unique Games conjecture, it is* **NP***-hard to approximate within a factor greater than* $27/32$ *the maximum revenue of a directed social network in the Uniform Additive Model.*

Proof. The proof is similar to the proof of Proposition 1. Let $G(V, E, w)$ be a directed social network, and let π^* be a vertex ordering corresponding to an acyclic subgraph of G with a maximum edge weight of W^*. Using π^* and a pricing probability of $2/3$ for all buyers, we obtain that the maximum revenue of G is at least $4W^*/27$.

We assume a marketing strategy $(\pi, \mathbf{p})$ that approximates the maximum revenue of G within a factor of r. Thus, $R(\pi, \mathbf{p}) \geq 4rW^*/27$. Let (i, j) be any edge with $\pi_i < \pi_j$, namely, any edge from which $(\pi, \mathbf{p})$ extracts some revenue. Since the revenue extracted from each such edge (i, j) is at most $w_{ij}/4$, the edge weight of the acyclic subgraph defined by π is at least $4R(\pi, \mathbf{p}) \geq \frac{16r}{27}W^*$

Thus, given an r-approximate marketing strategy $(\pi, \mathbf{p})$, we can approximate W^* within a ratio of $16r/27$. Now, the proposition follows from [64, Theorem 1.1]. □

5.2 Optimizing simple Influence and Exploit

A natural idea is to exploit the apparent connection between a large cut in the social network and a good IE strategy. For example, in the undirected case, an IE strategy $\mathrm{IE}(q, p)$ is conceptually similar to the randomized 0.5-approximation algorithm for MAX-CUT, which puts each node in set A with probability $1/2$. However, in addition to a revenue of $p(1-p)w_{ij}$ from each edge $\{i, j\}$ in the cut $(A, V \setminus A)$, $\mathrm{IE}(q, p)$ extracts a revenue of $p^2(1-p)w_{ij}$ from each edge $\{i, j\}$ between nodes in the exploit set $V \setminus A$. Thus, to optimize the performance

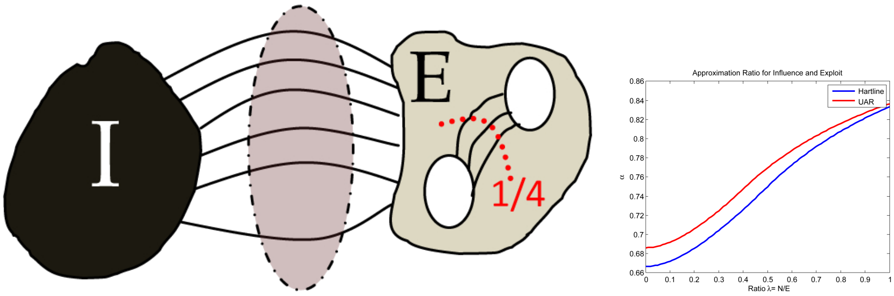


Figure 5.1: (a)Illustration of the simple IE strategy and (b) the approximation ratio as a function of the ration N/W

of $\mathrm{IE}(q,p)$, we carefully adjust the probabilities q and p so that $\mathrm{IE}(q,p)$ balances between the two sources of revenue. Hence, we obtain the following:

Proposition 3. *Let $G(V,E,w)$ be an undirected social network, and let $q = \max\{1-\frac{\sqrt{2}(2+\lambda)}{4}, 0\}$, where $\lambda = N/W$. Then, $\mathrm{IE}(q, 2-\sqrt{2})$ approximates the maximum revenue extracted from G within a factor of at least $2\sqrt{2}(2-\sqrt{2})(\sqrt{2}-1) \approx 0.686$.*

Proof. The proof extends the proof of [66, Theorem 3.1]. We start with calculating the expected (wrt to the random choice of the influence set) revenue of $\mathrm{IE}(q,p)$. The expected revenue of $\mathrm{IE}(q,p)$ from each loop $\{i,i\}$ is $(1-q)p(1-p)w_{ii}$. In particular, a revenue of $p(1-p)w_{ii}$ is extracted from $\{i,i\}$ if buyer i is included in the exploit set, which happens with probability $1-q$. The expected revenue of $\mathrm{IE}(q,p)$ from each edge $\{i,j\}$, $i<j$, is $(2q(1-q)p(1-p) + (1-q)^2p^2(1-p))w_{ij}$. More specifically, if one of i, j is included in the influence set and the other is included in the exploit set, which happens with probability $2q(1-q)$, a revenue of $p(1-p)w_{ij}$ is extracted from edge $\{i,j\}$. Otherwise, if both i and j are included in the exploit set, which happens with probability $(1-q)^2$, a revenue of $p^2(1-p)w_{ij}$ is extracted from edge $\{i,j\}$ (note that since $\{i,j\}$ is an undirected edge, the order in which i and j are considered in the exploit set is insignificant). By linearity of expectation, the expected revenue of $\mathrm{IE}(q,p)$ is:

$$R_{\mathrm{IE}}(q,p) = (1-q)p(1-p)\sum_{i\in V} w_{ii} + (1-q)p(1-p)\sum_{i<j}(2q+p(1-q))w_{ij}$$

Using that $N = \sum_{i\in V} w_{ii}$ and $W = \sum_{i<j} w_{ij}$, and setting $N = \lambda W$, we obtain that:

$$R_{\mathrm{IE}}(q,p) = (1-q)p(1-p)(\lambda + 2q + p(1-q))W$$

Differentiating with respect to q, we obtain that the optimal value of q is

$$q^* = \max\left\{\frac{1-p-\lambda/2}{2-p}, 0\right\}$$

We recall that $R^* = (1+\lambda)W/4$ is an upper bound on the maximum revenue of G. Therefore, the approximation ratio of $\mathrm{IE}(q,p)$ is:

$$\frac{4(1-q)p(1-p)(\lambda+2q+p(1-q))}{1+\lambda} \tag{5.3}$$

Using $p = 1/2$ and $q = \max\left\{\frac{1-\lambda}{3}, 0\right\}$ in (5.3), we obtain the IE strategy of [66, Theorem 3.1], whose approximation ratio is at least $2/3$, attained at $\lambda = 0$. Assuming small values of λ, so that $q^* > 0$, and differentiating with respect to p, we obtain that the best value of p for $\mathrm{IE}(q^*, p)$ is $p^* = 2 - \sqrt{2}$. Using $p = 2 - \sqrt{2}$ and $q = \max\left\{1 - \frac{\sqrt{2}(2+\lambda)}{4}, 0\right\}$, we obtain an IE strategy with an approximation ratio of at least $2\sqrt{2}(2-\sqrt{2})(\sqrt{2}-1) \approx 0.686$, attained at $\lambda = 0$. □

Proposition 4. *Let $G(V, E, w)$ be a directed social network. Then, $\mathrm{IE}\left(1 - \frac{\sqrt{2}}{2}, 2 - \sqrt{2}\right)$ approximates the maximum revenue of G within a factor of $\sqrt{2}(2 - \sqrt{2})(\sqrt{2} - 1) \approx 0.343$.*

Proof. The proof is similar to the proof of Proposition 3. We recall that for the directed case, we can ignore loops (i, i). Since the social network G is directed, the expected (wrt to the random choice of the influence set and the random order of the exploit set) revenue of $\mathrm{IE}(q, p)$ is:

$$\begin{aligned} R_{\mathrm{IE}}(q,p) &= (1-q)p(1-p)\sum_{(i,j)\in E}(q + p(1-q)/2)w_{ij} \\ &= (1-q)p(1-p)(q + p(1-q)/2)W \end{aligned}$$

More specifically, if i is included in the influence set and j is included in the exploit set, which happens with probability $q(1-q)$, a revenue of $p(1-p)w_{ij}$ is extracted from each edge (i, j). Furthermore, if both i and j are included in the exploit set $V \setminus A$ and i appears before j in the random order of $V \setminus A$, which happens with probability $(1-q)^2/2$, a revenue of $p^2(1-p)w_{ij}$ is extracted from edge (i, j).

Using the upper bound of $W/4$ on the maximum revenue of G, we have that the approximation ratio of $\mathrm{IE}(q, p)$ is at least $4(1-q)p(1-p)(q + p(1-q)/2)$. Setting $q = 1/3$ and $p = 1/2$, we obtain the IE strategy of [66, Theorem 3.1], whose approximation ratio for directed networks is $1/3$. Using $q = 1 - \frac{\sqrt{2}}{2}$ and $p = 2 - \sqrt{2}$, we obtain an IE strategy with an approximation ratio of $\sqrt{2}(2-\sqrt{2})(\sqrt{2}-1) \approx 0.343$. □

Proposition 5 (Optimality of IE for Bipartite Networks)**.** *Let $G(V, E, w)$ be an undirected bipartite social network with $w_{ii} = 0$ for all buyers i, and let $(A, V \setminus A)$ be any partition of V into independent sets. Then, $\mathrm{IE}(A, 1/2)$ extracts the maximum revenue of G.*

Proof. Since all edges of G are between buyers in the influence set A and buyers in the exploit set $V \setminus A$, $\mathrm{IE}(A, 1/2)$ extracts the myopic revenue of $w_{ij}/4$ from any edge $\{i, j\} \in E$. Therefore, $\mathrm{IE}(A, 1/2)$ is an optimal strategy. □

5.3 Generalized Influence and Exploit

Building on the idea of generating revenue from large cuts between different pricing classes, we obtain a class of generalized IE strategies, which employ a refined partition of buyers in more than two pricing classes. We first analyze the efficiency of generalized IE strategies for undirected networks, and then translate our results to the directed case. The analysis generalizes the proof of Proposition 3.

A generalized IE strategy consists of K pricing classes, for some appropriately large integer $K \geq 2$. Each class k, $k = 1, \ldots, K$, is associated with a pricing probability of $p_k = 1 - \frac{k-1}{2(K-1)}$. Each buyer is assigned to the pricing class k independently with probability q_k, where $\sum_{k=1}^{K} q_k = 1$, and is offered a pricing probability of p_k. The buyers are considered in non-increasing order of their pricing probabilities, i.e., the buyers in class k are considered before the buyers in class $k+1$, $k = 1, \ldots, K-1$. The buyers in the same class are considered in random order. In the following, we let $\mathrm{IE}(\mathbf{q}, \mathbf{p})$ denote such a generalized IE strategy, where $\mathbf{q} = (q_1, \ldots, q_K)$ is the assignment probability vector and $\mathbf{p} = (p_1, \ldots, p_K)$ is the pricing probability vector.

We proceed to calculate the expected revenue extracted by the generalized IE strategy $\mathrm{IE}(\mathbf{q}, \mathbf{p})$ from an undirected social network $G(V, E, w)$. The expected revenue of $\mathrm{IE}(\mathbf{q}, \mathbf{p})$ from each loop $\{i, i\}$ is $w_{ii} \sum_{k=1}^{K} q_k p_k (1 - p_k)$. Specifically, for each k, buyer i is included in the pricing class k with probability q_k, in which case, the revenue extracted from $\{i, i\}$ is $p_k(1 - p_k) w_{ii}$. The expected revenue of $\mathrm{IE}(\mathbf{p}, \mathbf{q})$ from each edge $\{i, j\}$, $i < j$, is:

$$w_{ij} \sum_{k=1}^{K} q_k p_k (1 - p_k) \left(q_k p_k + 2 \sum_{\ell=1}^{k-1} q_\ell p_\ell \right)$$

More specifically, for each class k, if both i, j are included in the pricing class k, which happens with probability q_k^2, the revenue extracted from $\{i, j\}$ is $p_k^2(1 - p_k) w_{ij}$. Furthermore, for each pair ℓ, k of pricing classes, $1 \leq \ell < k \leq K$, if either i is included in ℓ and j is included in k or the other way around, which happens with probability $2 q_\ell q_k$, the revenue extracted from $\{i, j\}$ is $p_\ell p_k (1 - p_k) w_{ij}$. Using linearity of expectation and setting $N = \sum_{i \in V} w_{ii}$ and $W = \sum_{i<j} w_{ij}$, we obtain that the expected revenue of $\mathrm{IE}(\mathbf{q}, \mathbf{p})$ is:

$$R_{\mathrm{IE}}(\mathbf{q}, \mathbf{p}) = N \sum_{k=1}^{K} q_k p_k (1 - p_k) + W \sum_{k=1}^{K} q_k p_k (1 - p_k) \left(q_k p_k + 2 \sum_{\ell=1}^{k-1} q_\ell p_\ell \right)$$

Since $R^* = (N + W)/4$ is an upper bound on the maximum revenue of G, the approximation ratio of $\mathrm{IE}(\mathbf{q}, \mathbf{p})$ is at least:

$$\min\left\{ 4 \sum_{k=1}^{K} q_k p_k (1 - p_k),\ \ 4 \sum_{k=1}^{K} q_k p_k (1 - p_k) \left(q_k p_k + 2 \sum_{\ell=1}^{k-1} q_\ell p_\ell \right) \right\} \tag{5.4}$$

We can now select the assignment probability vector $\mathbf{q}$ so that (5.4) is maximized. We note that with the pricing probability vector $\mathbf{p}$ fixed, this involves maximizing a quadratic function of $\mathbf{q}$ over linear constraints. Thus, we obtain the following:

Theorem 8. *For any undirected social network G, the generalized IE strategy with $K = 6$ pricing classes and assignment probabilities $\mathbf{q} = (0.183, 0.075, 0.075, 0.175, 0.261, 0.231)$ approximates the maximum revenue of G within a factor of* 0.7032*.*

We note that the approximation ratio can be improved to 0.706 by considering more pricing classes. By the same approach, we show that for directed social networks, the approximation ratio of $\mathrm{IE}(\mathbf{q}, \mathbf{p})$ is at least half the quantity in (5.4). Therefore:

Corollary 2. *For any directed social network G, the generalized IE strategy with $K = 6$ pricing classes and assignment probabilities $\mathbf{q} = (0.183, 0.075, 0.075, 0.175, 0.261, 0.231)$ approximates the maximum revenue of G within a factor of* 0.3516*.*

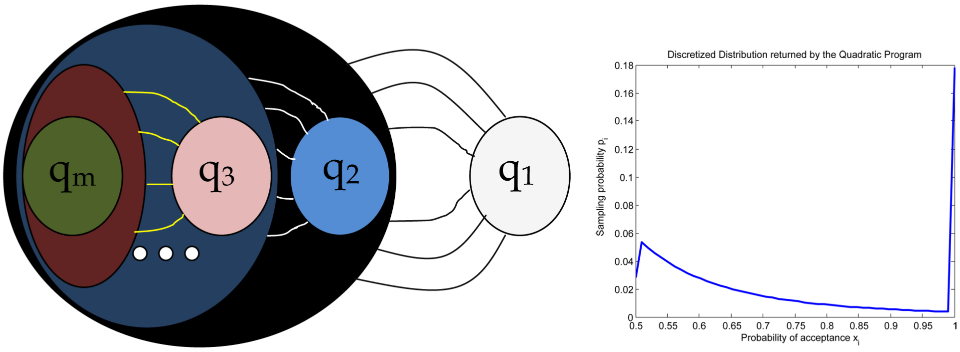


Figure 5.2: (a)Illustration of the generalized IE strategy and (b)the distribution over pricing classes returned by the quadratic program.

Proof. Similarly to the proof of Theorem 8, we calculate the expected (wrt the random partition of buyers into pricing classes and the random order of buyers in the pricing classes) revenue extracted by the generalized IE strategy $\text{IE}(\mathbf{p}, \mathbf{q})$ from a directed social network $G(V, E, w)$. We recall that for directed social networks, we can ignore loops (i, i). The expected revenue of $\text{IE}(\mathbf{p}, \mathbf{q})$ from each edge (i, j) is:

$$w_{ij} \sum_{k=1}^{K} q_k p_k (1 - p_k) \left(\frac{q_k p_k}{2} + \sum_{\ell=1}^{k-1} q_\ell p_\ell \right)$$

More specifically, for each class k, if both i, j are included in the pricing class k and i appears before j in the random order of the buyers in k, which happens with probability $q_k^2/2$, the revenue extracted from each edge (i, j) is $p_k^2(1 - p_k)w_{ij}$. Furthermore, for each pair ℓ, k of pricing classes, $1 \le \ell < k \le K$, if i is included in ℓ and j is included in k, which happens with probability $q_\ell q_k$, the revenue extracted from (i, j) is $p_\ell p_k (1 - p_k) w_{ij}$.

Using linearity of expectation and setting $W = \sum_{(i,j) \in E} w_{ij}$, we obtain that the expected revenue of $\text{IE}(\mathbf{q}, \mathbf{p})$ is:

$$R_{\text{IE}}(\mathbf{q}, \mathbf{p}) = W \sum_{k=1}^{K} q_k p_k (1 - p_k) \left(\frac{q_k p_k}{2} + \sum_{\ell=1}^{k-1} q_\ell p_\ell \right)$$

Since $W/4$ is an upper bound on the maximum revenue of G, the approximation ratio of $\text{IE}(\mathbf{q}, \mathbf{p})$ is at least:

$$4 \sum_{k=1}^{K} q_k p_k (1 - p_k) \left(\frac{q_k p_k}{2} + \sum_{\ell=1}^{k-1} q_\ell p_\ell \right) , \qquad (5.5)$$

namely at least half of the approximation ratio in the undirected case.

Using $\mathbf{q} = (0.183, 0.075, 0.075, 0.175, 0.261, 0.231)$ in (5.5), we obtain an approximation ratio of at least 0.3516. □

5.4 Influence and Exploit via Semidefinite Programming

The main hurdle in obtaining better approximation guarantees for the maximum revenue problem is the loose upper bound of $(N + W)/4$ on the optimal revenue. We do not know how to obtain a stronger upper bound on the maximum revenue. However, in this section, we obtain a strong Semidefinite Programming (SDP) relaxation for the problem of computing the best IE strategy with any given pricing probability $p \in [1/2, 1)$. Our approach exploits the resemblance between computing the best IE strategy and the problems of MAX-CUT (for undirected networks) and MAX-DICUT (for directed networks), and builds on the elegant approach of Goemans and Williamson [55] and Feige and Goemans [45]. Solving the SDP relaxation and using randomized rounding, we obtain, in polynomial time, a good approximation to the best influence set for the given pricing probability p. Then, employing the bounds of Theorem 6 and Theorem 7, we obtain strong approximation guarantees for the maximum revenue problem for both directed and undirected networks. The high level description of our algorithm **SDP-IE** is:

SDP-IE

Input: *A weighted directed graph $G(V, E)$and a number ϵ*
SDP-relaxation:
1. *Solve the Semi-Definite relaxation (**??**) with accuracy $(1-\epsilon)$ to obtain vectors v_i, $i = 0, 1, \ldots, n$.*
Rotation:
2. *Obtain the rotated vectors r_i, $i = 1, \ldots, n$, where the rotation is made by a function $f_\lambda(\theta)$:*

$$f_\lambda(\theta) = (1-\lambda)\theta + \lambda\frac{\pi}{2}(1-\cos\theta) \tag{5.6}$$

Randomized Rounding:
3. *Select a random vector $r \in S^n$.*
4. ***if*** *$sign(r_i \cdot r) = sign(v_0 \cdot r)$ put vertex i in the Influence set*
else *put it in the Exploit set.*
Output: *the Influence and Exploit sets.*

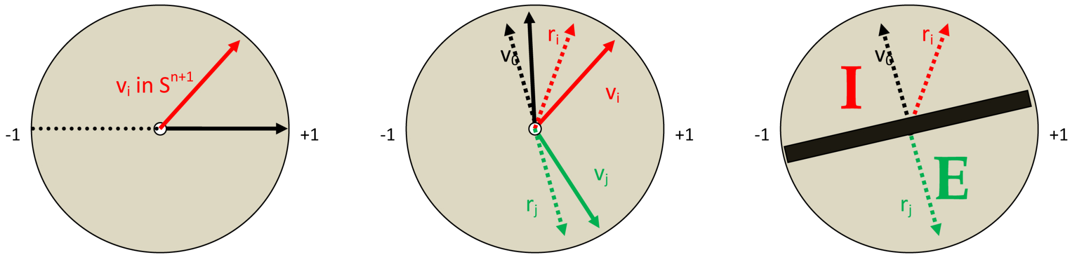
Since Influence and Exploit requires a binary decision to be made for each vertex, the general idea behind this approach is to use randomization. That is, to construct an appropriate probability space(assigning probabilities to each possible Influence and Exploit set pair) in which the expectation of the revenue is high. Thus, by solving the semidefinite relaxation we obtain a favorable underlying geometry through which the randomized decision rule defines the required probability space. The rotation of the vectors is an intermediate step which increases the probability that heavy edges are cut. We proceed with the analysis of the algorithm.

Directed Social Networks. We start with the case of a directed social network $G(V, E, w)$, which is a bit simpler, because we can ignore loops (i, i) without loss of generality. We observe that for any given pricing probability $p \in [1/2, 1)$, the problem of computing the best IE strategy IE(A, p) is equivalent to solving the following Quadratic Integer Program:

$$\max \tfrac{p(1-p)}{4} \sum_{(i,j)\in E} w_{ij} \left(1 + \tfrac{p}{2} + (1-\tfrac{p}{2})y_0 y_i - (1+\tfrac{p}{2})y_0 y_j - (1-\tfrac{p}{2})y_i y_j\right) \tag{Q1}$$

$$\text{s.t.} \qquad y_i \in \{-1, 1\} \qquad \forall i \in V \cup \{0\}$$

In (Q1), there is a variable y_i for each buyer i and an additional variable y_0 denoting the influence set. A buyer i is assigned to the influence set A, if $y_i = y_0$, and to the exploit



(a) Relaxation (b) Rotation (c) Hyperplane Rounding

Figure 5.3: Graphical depiction of the main steps of the SDP-IEalgorithm.

set, otherwise. For each edge (i, j), $1 + y_0y_i - y_0y_j - y_iy_j$ is 4, if $y_i = y_0 = -y_j$ (i.e., if i is assigned to the influence set and j is assigned to the exploit set), and 0, otherwise. Moreover, $\frac{p}{2}(1 - y_0y_i - y_0y_j + y_iy_j)$ is $2p$, if $y_i = y_j = -y_0$ (i.e., if both i and j are assigned to the exploit set), and 0, otherwise. Therefore, the contribution of each edge (i, j) to the objective function of (Q1) is equal to the revenue extracted from (i, j) by IE(A, p).

Following the approach of [55, 45], we relax (Q1) to the following Semidefinite Program, where $v_i \cdot v_j$ denotes the inner product of vectors v_i and v_j:

$$\max \frac{p(1-p)}{4} \sum_{(i,j)\in E} w_{ij} \left(1 + \tfrac{p}{2} + (1 - \tfrac{p}{2})\, v_0 \cdot v_i - (1 + \tfrac{p}{2})\, v_0 \cdot v_j - (1 - \tfrac{p}{2})\, v_i \cdot v_j\right) \qquad \text{(S1)}$$

$$\begin{aligned} \text{s.t.} \qquad & v_i \cdot v_j + v_0 \cdot v_i + v_0 \cdot v_j \geq -1 \\ & v_i \cdot v_j - v_0 \cdot v_i - v_0 \cdot v_j \geq -1 \\ & -v_i \cdot v_j - v_0 \cdot v_i + v_0 \cdot v_j \geq -1 \\ & -v_i \cdot v_j + v_0 \cdot v_i - v_0 \cdot v_j \geq -1 \\ & v_i \cdot v_i = 1, \quad v_i \in \mathbb{R}^{n+1} \qquad \forall i \in V \cup \{0\} \end{aligned}$$

We observe that any feasible solution to (Q1) can be translated into a feasible solution to (S1) by setting $v_i = v_0$, if $y_i = y_0$, and $v_i = -v_0$, otherwise. An optimal solution to (S1) can be computed within any precision ε in time polynomial in n and in $\ln \frac{1}{\varepsilon}$ (see e.g. [3]).

Given a directed social network $G(V, E, w)$, a pricing probability p, and a parameter $\gamma \in [0, 1]$, the algorithm SDP-IE(p, γ) first computes an optimal solution $v_0, v_1, \ldots, v_n$ to (S1). Then, following [45], the algorithm maps each vector v_i to a rotated vector v'_i which is coplanar with v_0 and v_i, lies on the same side of v_0 as v_i, and forms an angle with v_0 equal to

$$f_\gamma(\theta_i) = (1 - \gamma)\theta_i + \gamma\pi(1 - \cos\theta_i)/2\,,$$

where $\pi = 3.14\ldots$ and $\theta_i = \arccos(v_0 \cdot v_i)$ is the angle of v_0 and v_i. Finally, the algorithm computes a random vector r uniformly distributed on the unit $(n + 1)$-sphere, and assigns each buyer i to the influence set A, if $\text{sgn}(v'_i \cdot r) = \text{sgn}(v_0 \cdot r)$, and to the exploit set $V \setminus A$, otherwise[1], where $\text{sgn}(x) = 1$, if $x \geq 0$, and -1, otherwise. We next show that:

[1]Let $\theta'_i = \arccos(v_0 \cdot v'_i)$ be the angle of v_0 and a rotated vector v'_i. To provide some intuition behind the rotation step, we note that $\theta'_i < \theta_i$, if $\theta_i \in (0, \pi/2)$, and $\theta'_i > \theta_i$, if $\theta_i \in (\pi/2, \pi)$. Therefore, applying rotation to v_i, the algorithm increases the probability of assigning i to the influence set, if $\theta_i \in (0, \pi/2)$, and the probability of assigning i to the exploit set, if $\theta_i \in (\pi/2, \pi)$. The strength of the rotation's effect depends on the value of γ and on the value of θ_i.

Theorem 9. *For any directed social network G, SDP-IE$(2/3, 0.722)$ approximates the maximum revenue extracted from G by the best IE strategy with pricing probability $2/3$ within a factor of* 0.9064.

Proof. In the following, we let $v_0, v_1, \ldots, v_n$ be an optimal solution to (S1), let $\theta_{ij} = \arccos(v_i \cdot v_j)$ be the angle of any two vectors v_i and v_j, and let $\theta_i = \arccos(v_0 \cdot v_i)$ be the angle of v_0 and any vector v_i. Similarly, we let $\theta'_{ij} = \arccos(v'_i \cdot v'_j)$ be the angle of any two rotated vectors v'_i and v'_j, and let $\theta'_i = \arccos(v_0 \cdot v'_i)$ be the angle of v_0 and any rotated vector v'_i. We first calculate the expected revenue extracted from each edge $(i, j) \in E$ by the IE strategy of SDP-IE(p, γ).

Lemma 8. *The IE strategy of SDP-IE(p, γ) extracts from each edge (i, j) an expected revenue of:*

$$w_{ij}\, p(1-p)\, \frac{(1-\frac{p}{2})\,\theta'_{ij} - (1-\frac{p}{2})\,\theta'_i + (1+\frac{p}{2})\,\theta'_j}{2\pi} \tag{5.7}$$

Proof. We first define the following mutually disjoint events:

$$\begin{aligned}
B^{ij} &: \operatorname{sgn}(v'_i \cdot r) = \operatorname{sgn}(v'_j \cdot r) = \operatorname{sgn}(v_0 \cdot r) \\
B^i_j &: \operatorname{sgn}(v'_i \cdot r) = \operatorname{sgn}(v_0 \cdot r) \neq \operatorname{sgn}(v'_j \cdot r) \\
B^j_i &: \operatorname{sgn}(v'_j \cdot r) = \operatorname{sgn}(v_0 \cdot r) \neq \operatorname{sgn}(v'_i \cdot r) \\
B_{ij} &: \operatorname{sgn}(v'_i \cdot r) = \operatorname{sgn}(v'_j \cdot r) \neq \operatorname{sgn}(v_0 \cdot r)
\end{aligned}$$

Namely, B^{ij} (resp. B_{ij}) is the event that both i and j are assigned to the influence set A (resp. to the exploit set $V \setminus A$), and B^i_j (resp. B^j_i) is the event that i (resp. j) is assigned to the influence set A and j (resp. i) is assigned to the exploit set $V \setminus A$. Also, we let $\mathbb{P}\mathrm{r}[B]$ denote the probability of any event B. Then, the expected revenue extracted from each edge (i, j) is:

$$w_{ij}\, p(1-p)\, \left(\mathbb{P}\mathrm{r}[B^i_j] + \tfrac{p}{2}\, \mathbb{P}\mathrm{r}[B_{ij}]\right) \tag{5.8}$$

To calculate $\mathbb{P}\mathrm{r}[B^i_j]$ and $\mathbb{P}\mathrm{r}[B_{ij}]$, we use that if r is a vector uniformly distributed on the unit sphere, for any vectors v_i, v_j on the unit sphere, $\mathbb{P}\mathrm{r}[\operatorname{sgn}(v_i \cdot r) \neq \operatorname{sgn}(v_j \cdot r)] = \theta_{ij}/\pi$ [55, Lemma 3.2]. For $\mathbb{P}\mathrm{r}[B^i_j]$, we calculate the probability of the event $B^i_j \cup B^j_i$ that i and j are in different sets, of the event $B^i_j \cup B^{ij}$ that i is in the influence set, and of the event $B^j_i \cup B^{ij}$ that j is in the influence set.

$$\mathbb{P}\mathrm{r}[B^i_j] + \mathbb{P}\mathrm{r}[B^j_i] = \mathbb{P}\mathrm{r}[B^i_j \cup B^j_i] = \mathbb{P}\mathrm{r}[\operatorname{sgn}(v'_i \cdot r) \neq \operatorname{sgn}(v'_j \cdot r)] = \theta'_{ij}/\pi \tag{5.9}$$

$$\mathbb{P}\mathrm{r}[B^i_j] + \mathbb{P}\mathrm{r}[B^{ij}] = \mathbb{P}\mathrm{r}[B^i_j \cup B^{ij}] = \mathbb{P}\mathrm{r}[\operatorname{sgn}(v'_i \cdot r) = \operatorname{sgn}(v_0 \cdot r)] = 1 - \theta'_i/\pi \tag{5.10}$$

$$\mathbb{P}\mathrm{r}[B^j_i] + \mathbb{P}\mathrm{r}[B^{ij}] = \mathbb{P}\mathrm{r}[B^j_i \cup B^{ij}] = \mathbb{P}\mathrm{r}[\operatorname{sgn}(v'_j \cdot r) = \operatorname{sgn}(v_0 \cdot r)] = 1 - \theta'_j/\pi \tag{5.11}$$

Subtracting (5.11) from (5.9) plus (5.10), we obtain that:

$$\mathbb{P}\mathrm{r}[B^i_j] = \tfrac{1}{2\pi}(\theta'_{ij} - \theta'_i + \theta'_j) \tag{5.12}$$

For $\mathbb{P}\mathrm{r}[B_{ij}]$, we also need the probability of the event $B^j_i \cup B_{ij}$ that i is in the exploit set, and of the event $B^i_j \cup B_{ij}$ that j is in the exploit set.

$$\mathbb{P}\mathrm{r}[B^j_i] + \mathbb{P}\mathrm{r}[B_{ij}] = \mathbb{P}\mathrm{r}[B^j_i \cup B_{ij}] = \mathbb{P}\mathrm{r}[\operatorname{sgn}(v'_i \cdot r) \neq \operatorname{sgn}(v_0 \cdot r)] = \theta'_i/\pi \tag{5.13}$$

$$\mathbb{P}\mathrm{r}[B^i_j] + \mathbb{P}\mathrm{r}[B_{ij}] = \mathbb{P}\mathrm{r}[B^i_j \cup B_{ij}] = \mathbb{P}\mathrm{r}[\operatorname{sgn}(v'_j \cdot r) \neq \operatorname{sgn}(v_0 \cdot r)] = \theta'_j/\pi \tag{5.14}$$

Subtracting (5.9) from (5.13) plus (5.14), we obtain that:

$$\mathbb{P}\mathrm{r}[B_{ij}] = \tfrac{1}{2\pi}(-\theta'_{ij} + \theta'_i + \theta'_j) \tag{5.15}$$

Substituting (5.12) and (5.15) in (5.8), we obtain (5.7), and conclude the proof of the lemma. □

Since (S1) is a relaxation of the problem of computing the best IE strategy with pricing probability p, the revenue of an optimal IE(A, p) strategy is at most:

$$\tfrac{p(1-p)}{4} \sum_{(i,j)\in E} w_{ij} \left(1 + \tfrac{p}{2} + (1 - \tfrac{p}{2}) \cos\theta_i - (1 + \tfrac{p}{2}) \cos\theta_j - (1 - \tfrac{p}{2}) \cos\theta_{ij}\right) \tag{5.16}$$

On the other hand, by Lemma 8 and linearity of expectation, the IE strategy of SDP-IE(p, γ) generates an expected revenue of:

$$\tfrac{p(1-p)}{2\pi} \sum_{(i,j)\in E} w_{ij} \left((1 - \tfrac{p}{2})\, \theta'_{ij} - (1 - \tfrac{p}{2})\, \theta'_i + (1 + \tfrac{p}{2})\, \theta'_j\right) \tag{5.17}$$

We recall that for each i, $\theta'_i = f_\gamma(\theta_i)$. Moreover, in [45, Section 4], it is shown that for each i, j,

$$\theta'_{ij} = g_\gamma(\theta_{ij}, \theta_i, \theta_j) = \arccos\left(\cos f_\gamma(\theta_i)\, \cos f_\gamma(\theta_j) + \frac{\cos\theta_{ij} - \cos\theta_i\, \cos\theta_j}{\sin\theta_i\, \sin\theta_j}\, \sin f_\gamma(\theta_i)\, \sin f_\gamma(\theta_j)\right)$$

The approximation ratio of SDP-IE(p, γ) is derived as the minimum ratio of any pair of terms in (5.17) and (5.16) corresponding to the same edge (i, j). Thus, the approximation ratio of SDP-IE(p, γ) is:

$$\begin{aligned}
\rho(p, \gamma) = \frac{2}{\pi} \min_{0\le x,y,z\le\pi} & \frac{(1 - \frac{p}{2})\, g_\gamma(x, y, z) - (1 - \frac{p}{2}) f_\gamma(y) + (1 + \frac{p}{2}) f_\gamma(z)}{1 + \frac{p}{2} + (1 - \frac{p}{2}) \cos y - (1 + \frac{p}{2}) \cos z - (1 - \frac{p}{2}) \cos x} \\
\text{s.t.} \qquad & \cos x + \cos y + \cos z \ge -1 \\
& \cos x - \cos y - \cos z \ge -1 \\
& -\cos x - \cos y + \cos z \ge -1 \\
& -\cos x + \cos y - \cos z \ge -1
\end{aligned}$$

It can be shown numerically, that $\rho(2/3, 0.722) \ge 0.9064$. □

Combining Theorem 9 and Theorem 7, we conclude that:

Theorem 10. *For any directed social network G, the IE strategy computed by SDP-IE$(2/3, 0.722)$ approximates the maximum revenue of G within a factor of* 0.5011.

Undirected Social Networks. We apply the same approach to an undirected network $G(V, E, w)$. For any given pricing probability $p \in [1/2, 1)$, the problem of computing the best IE strategy IE(A, p) for G is equivalent to solving the following Quadratic Integer Program:

$$\begin{aligned}
\max\ & \tfrac{p(1-p)}{2} \sum_{i\in V} w_{ii}\, (1 - y_0 y_i) + \\
+\ & \tfrac{p(1-p)}{4} \sum_{i<j} w_{ij}\, (2 + p - p y_0 y_i - p y_0 y_j - (2 - p) y_i y_j)
\end{aligned} \tag{Q2}$$

$$\text{s.t.} \qquad y_i \in \{-1, 1\} \qquad \forall i \in V \cup \{0\}$$

In (Q2), there is a variable y_i for each buyer i and an additional variable y_0 denoting the influence set. A buyer i is assigned to the influence set A, if $y_i = y_0$, and to the exploit set, otherwise. For each loop $\{i, i\}$, $1 - y_0 y_i$ is 2, if i is assigned to the exploit set, and 0, otherwise. For each edge $\{i, j\}$, $i < j$, $2 - 2y_i y_j$ is 4, if i and j are assigned to different sets, and 0, otherwise. Also, $p(1 - y_0 y_i - y_0 y_j + y_i y_j)$ is $4p$, if both i and j are assigned to the exploit set, and 0, otherwise. Therefore, the contribution of each loop $\{i, i\}$ and each edge $\{i, j\}$, $i < j$, to the objective function of (Q2) is equal to the revenue extracted from them by IE(A, p). The next step is to relax (Q1) to the following Semidefinite Program:

$$\begin{aligned}
\max \quad & \tfrac{p(1-p)}{2} \textstyle\sum_{i \in V} w_{ii} \, (1 - v_0 \cdot v_i) + \\
+ \quad & \tfrac{p(1-p)}{4} \textstyle\sum_{i<j} w_{ij} \, (2 + p - p \, v_0 \cdot v_i \quad - p \, v_0 \cdot v_j - (2-p) \, v_i \cdot v_j) \\
\text{s.t.} \quad & v_i \cdot v_j + v_0 \cdot v_i + v_0 \cdot v_j \geq -1 \\
& v_i \cdot v_j - v_0 \cdot v_i - v_0 \cdot v_j \geq -1 \\
& -v_i \cdot v_j - v_0 \cdot v_i + v_0 \cdot v_j \geq -1 \\
& -v_i \cdot v_j + v_0 \cdot v_i - v_0 \cdot v_j \geq -1 \\
& v_i \cdot v_i = 1, \quad v_i \in \mathbb{R}^{n+1} \qquad \forall i \in V \cup \{0\}
\end{aligned} \tag{S2}$$

The algorithm is the same as the algorithm for directed networks. Specifically, given an undirected social network $G(V, E, w)$, a pricing probability p, and a parameter $\gamma \in [0, 1]$, the algorithm SDP-IE(p, γ) first computes an optimal solution $v_0, v_1, \ldots, v_n$ to (S2). Then, it maps each vector v_i to a rotated vector v'_i which is coplanar with v_0 and v_i, lies on the same side of v_0 as v_i, and forms an angle $f_\gamma(\theta_i)$ with v_0, where $\theta_i = \arccos(v_0 \cdot v_i)$. Finally, the algorithm computes a random vector r uniformly distributed on the unit $(n+1)$-sphere, and assigns each buyer i to the influence set A, if $\mathrm{sgn}(v'_i \cdot r) = \mathrm{sgn}(v_0 \cdot r)$, and to the exploit set $V \setminus A$, otherwise. We prove that:

Theorem 11. *For any undirected network G, SDP-IE$(0.586, 0.209)$ approximates the maximum revenue extracted from G by the best IE strategy with pricing probability 0.586 within a factor of 0.9032.*

Proof. We employ the same approach, techniques, and notation as in the proof of Theorem 9. The expected revenue extracted from each loop $\{i, i\}$ is $w_{ii} \, p(1-p)$ times the probability that i is in the exploit set, which is equal to $\mathbb{P}\mathrm{r}[\mathrm{sgn}(v'_i \cdot r) \neq \mathrm{sgn}(v_0 \cdot r)] = \theta'_i / \pi$. Therefore, the algorithm extracts an expected revenue of $w_{ii} \, p(1-p) \, \theta'_i / \pi$ from each loop $\{i, i\}$. Next, we calculate the expected revenue extracted from each (undirected) edge $\{i, j\}$, $i < j$, by the IE strategy of SDP-IE(p, γ).

Lemma 9. *SDP-IE(p, γ) extracts from each edge $\{i, j\}$, $i < j$, an expected revenue of:*

$$w_{ij} \, p(1-p) \, \frac{(2-p) \, \theta'_{ij} + p \, \theta'_i + p \, \theta'_j}{2\pi}$$

Proof. Let the events B^i_j, B^j_i, and B_{ij} be defined as in the proof of Lemma 8. In particular, $B^i_j \cup B^j_i$ is the event that i and j are in different sets, and B_{ij} is the event that both i and j are in the exploit set. Thus, the expected revenue extracted from edge $\{i, j\}$ is:

$$w_{ij}\, p(1-p) \left(\mathbb{P}\mathrm{r}[B^i_j \cup B^j_i] + p\, \mathbb{P}\mathrm{r}[B_{ij}] \right) \tag{5.18}$$

In the proof of Lemma 8, in (5.9) and (5.15) respectively, we show that $\mathbb{P}\mathrm{r}[B^i_j \cup B^j_i] = \theta'_{ij}/\pi$, and that $\mathbb{P}\mathrm{r}[B_{ij}] = (-\theta'_{ij}+\theta'_i+\theta'_j)/(2\pi)$. Substituting these in (5.18), we obtain the lemma. □

Therefore, by linearity of expectation, the expected revenue of SDP-IE(p, γ) is:

$$\tfrac{p(1-p)}{\pi} \sum_{i \in V} w_{ii}\, \theta'_i + \tfrac{p(1-p)}{2\pi} \sum_{i<j} w_{ij} \left((2-p)\, \theta'_{ij} + p\, \theta'_i + p\, \theta'_j \right) , \tag{5.19}$$

where $\theta'_i = f_\gamma(\theta_i)$, for each $i \in V$, and $\theta'_{ij} = g_\gamma(\theta_{ij}, \theta_i, \theta_j)$, for each $i, j \in V$.

On the other hand, since (S2) relaxes the problem of computing the best IE strategy with pricing probability p, the revenue of the best IE(A, p) strategy is at most:

$$\tfrac{p(1-p)}{2} \sum_{i \in V} w_{ii} (1 - \cos\theta_i) + \tfrac{p(1-p)}{4} \sum_{i<j} w_{ij} \left(2 + p - p\cos\theta_i - p\cos\theta_j - (2-p)\cos\theta_{ij} \right) \tag{5.20}$$

The approximation ratio of SDP-IE(p, γ) is derived as the minimum ratio of any pair of terms in (5.19) and (5.20) corresponding either to the same loop $\{i, i\}$ or to the same edge $\{i, j\}$, $i < j$. Therefore, the approximation ratio of SDP-IE(p, γ) for undirected social networks is the minimum of $\rho_1(\gamma)$ and $\rho_2(p, \gamma)$, where:

$$\begin{aligned}
\rho_1(\gamma) &= \frac{2}{\pi} \min_{0 \le x \le \pi} \frac{f_\gamma(x)}{1 - \cos x} \qquad \text{and} \\
\rho_2(p, \gamma) &= \frac{2}{\pi} \min_{0 \le x,y,z \le \pi} \frac{(2-p)\, g_\gamma(x,y,z) + p f_\gamma(y) + p f_\gamma(z)}{2 + p - p\cos y - p\cos z - (2-p)\cos x} \\
&\quad \text{s.t.} \quad \cos x + \cos y + \cos z \ge -1 \\
&\quad\quad\quad \cos x - \cos y - \cos z \ge -1 \\
&\quad\quad\quad -\cos x - \cos y + \cos z \ge -1 \\
&\quad\quad\quad -\cos x + \cos y - \cos z \ge -1
\end{aligned}$$

It can be shown numerically, that $\rho_1(0.209) \ge 0.9035$ and that $\rho_2(0.586, 0.209) \ge 0.9032$. □

Combining Theorem 11 and Theorem 6, we conclude that:

Theorem 12. *For any undirected social network G, the IE strategy computed by SDP-IE*$(0.586, 0.209)$ *approximates the maximum revenue of G within a factor of* 0.8229.

Remark. We can use $\rho(p, \gamma)$ and $\min\{\rho_1(\gamma), \rho_2(p, \gamma)\}$, and compute the approximation ratio of SDP-IE(p, γ) for the best IE strategy with any given pricing probability $p \in [1/2, 1)$. We note that $\rho_1(\gamma)$ is ≈ 0.87856, for $\gamma = 0$ (see e.g. [55, Lemma 3.5]), and increases slowly with γ. Viewed as a function of p, the value of γ maximizing $\rho(p, \gamma)$ and $\rho_2(p, \gamma)$ and the corresponding approximation ratio for the revenue of the best IE strategy increase slowly with p (see also Fig 5.4 about the dependence of γ and the approximation ratio as a function of p). For example, for directed social networks, the approximation ratio of SDP-IE($0.5, 0.653$) (resp. SDP-IE($0.52, 0.685$) and SDP-IE($0.52, 0.704$)) is 0.8942 (resp. 0.8955 and 0.9005). For undirected networks, the ratio of SDP-IE($0.5, 0.176$) (resp. SDP-IE($0.52, 0.183$) and SDP-IE($2/3, 0.425$)) is 0.899 (resp. 0.9005 and 0.907).

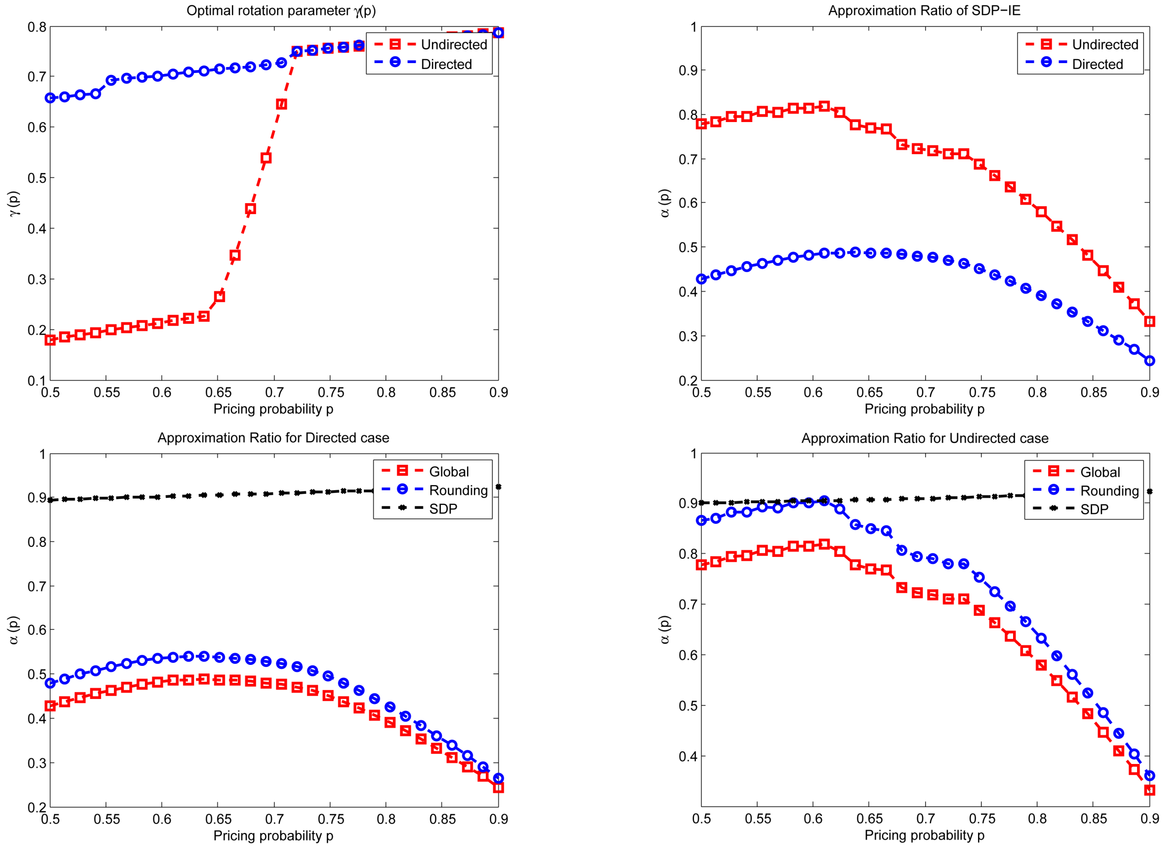


Figure 5.4: The approximation ratio of SDP-IE(p, γ) for the revenue of the best IE strategy and for the maximum revenue, as a function of the pricing probability p. The upper left plot shows the best choice of the rotation parameter γ, as a function of p. The blue curve (with circles) shows the best choice of γ for directed social networks and the red curve (with squares) for undirected networks. In both cases, the best choice of γ increases with p. The upper right plot shows the approximation ratio of SDP-IE(p, γ) for the maximum revenue for directed (blue curve, with circles) and undirected (red curve, with squares) networks. The lower plots show the approximation ratio of SDP-IE(p, γ) for directed (left plot) and undirected (right plot) networks, as a function of p. In each plot, the upper curve (in black) shows the approximation ratio of SDP-IE(p, γ) for the revenue of the best IE strategy, which increases slowly with p. The blue curve (that with circles) shows the guarantee of Theorem 7 and Theorem 6 on the fraction of the maximum revenue extracted by the best IE strategy. The red curve (that with squares) shows the approximation ratio of SDP-IE(p, γ) for the maximum revenue.

Chapter 6

Local Search and Heuristics

The Revenue Maximization problem is not only interesting from a theoretical point of view but from a practical as well. Specifically, a seller would want to improve a given marketing strategy if possible without needing any guarantee of how good the improvement would be. Motivated by this fact, in this section we introduce a class of Local Search strategies designed to improve a given marketing strategy $(\pi, \vec{p})$. We propose two special instantiations and discuss convergence issues. Furthermore, since the powerful SDP algorithms developed have large running time, their applications on massive graphs though interesting from a theoretical perspective becomes impractical. We address this issue by proposing intelligent heuristic based on eigenvector centrality[20, 83] correlating network position with price to be offered. Lastly, we discuss another approach on designing pricing strategies that generalize the Influence and Exploit idea to the furthest extent utilizing Calculus of Variations[24, 51].

We first provide some preliminary facts that will motivate later developments. A quantity that provides intuition is the *Forward Looking Revenue (FLR)*. Hartline et.al [66] considered the case of the optimal myopic revenue. The FLR is an insightful generalization. Consider that we are about to offer a price to buyer i, let A_i and B_i be the set of vertices that are considered after and before i respectively. The FLR consists of two parts; the expected revenue that we extract from buyer i and the extra revenue that becomes available if i accepts our offer, due to i's influence in the network:

$$FLR_i = \underbrace{p_i(1-p_i)\left(\sum_{j\in B_i} p_j w_{ji} + w_{ii}\right)}_{Revenue\ \ from\ \ i} + \underbrace{p_i \sum_{j\in A_i} p_j(1-p_j)w_{ij}}_{Extra\ \ Revenue\ \ if\ \ i\ \ accepts} \tag{6.1}$$

Viewed alternatively FLR_i is the part of the total revenue that depends on p_i.

Impact of Social Position. Assume for a moment that we are given both the sequence π and all the probabilities $\vec{p}_{-i}$ except for p_i. It turns out that in that case we can make an optimal choice (best response strategy) for p_i by maximizing FLR_i, while requiring that p_i is a valid probability between 0 and 1. Fortunately, when FLR_i is restricted only as a function of p_i it is a concave function and thus has a unique maximum for:

$$p_i^* = \left[\frac{1}{2} + \frac{1}{2}\frac{\sum_{j\in A_i} p_j(1-p_j)w_{ij}}{\sum_{j\in B_i} p_j w_{ji} + w_{ii}}\right]_{\leq 1} = \frac{1}{2} + \frac{1}{2}\left[\frac{N_i}{I_i}\right]_{\leq 1} \tag{6.2}$$

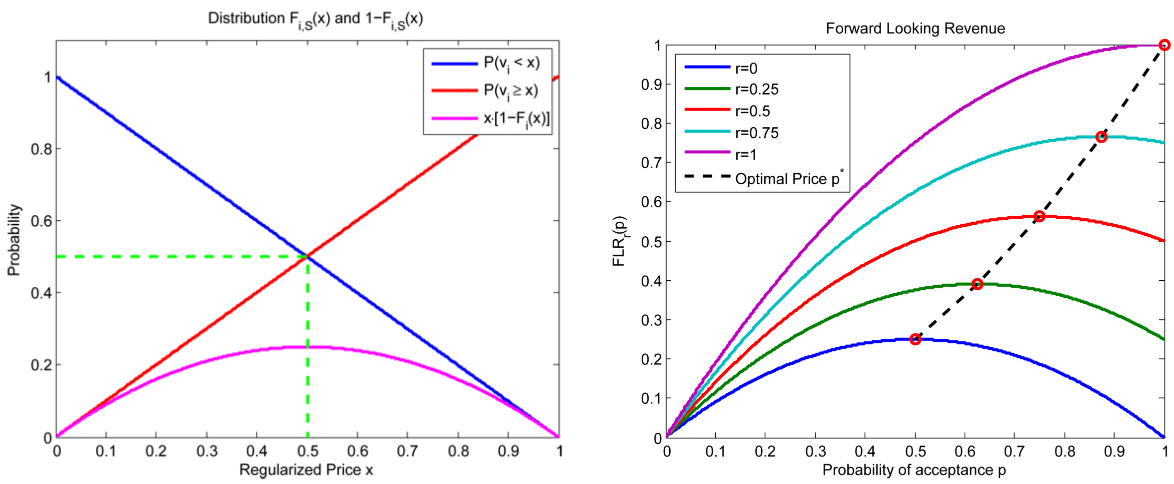


Figure 6.1: (a)Optimal myopic revenue(purple) as a function of the pricing probability x.(b)Forward looking revenue as a function of the pricing probability for different ratios $r = N_i/I_i$.

The above equation reveals a number of interesting facts. First, it states that all probabilities must lie in $[0.5, 1]$. Furthermore, observe that the "optimal" probability p_i consists of two terms:

1. The first term $1/2$ corresponds to the myopic price, i.e. the probability with which we would have i accept our offer disregarding network effects.

2. The second term is proportional to the ratio $r_i = N_i/I_i$ of the "network value" to the "intrinsic value" of node i. This term is essentially a *discount* offered to i in order to strike balance between the revenue we exert from him and the influence he has on the network.

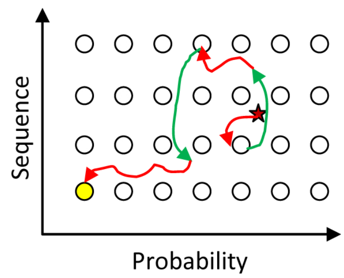
6.1 Local Search

Our general approach in designing these strategies is to utilize the optimality condition (6.2) along with Lemma 4(Chapter 4) for the ordering of buyers. These conditions show that there is a recursive dependence between ordering and probabilities that prevents us from computing optimally either one. In order to break this cycle of dependence we propose iterative schemes that after each step guarantee an increase in revenue. These strategies consist of two different functions that correspond to the dual nature (sequence - probabilities) of our problem and are reminiscent of a class of algorithms called *Expectation-Maximization*, which are extensively used in Machine Learning[17]. Particularly, these algorithms are compositions of two functions:

1. **Sequence Function**(*Expectation Step*): where we apply the sequence lemma in order to find a better ordering given the probability vector $\vec{p}$.

2. **Probability Function**(*Maximization Step*): where we iteratively maximize the *Forward Looking Revenue* for one variable at a time using some (cyclic) rule.

The increase in revenue for the first function is guaranteed by Lemma 4, whereas the concavity of FLR_i, when viewed only as a function of p_i, guarantees the increase for the second function. We proceed with some comments describing the nature of these algorithms.



Imagine that we enumerated all possible sequences X and all possible discretized probability vectors Y. If we take the cartesian product of these two sets, we would get a two dimensional grid $X \times Y$ with arbitrary connections between nodes. Particularly, all nodes with same x coordinate correspond to the different probability vector configurations given the sequence, whereas in the other case they correspond to the different sequences for a given probability vector. In this setting, our algorithms start from a node and walk on the grid making vertical and horizontal moves only, improving each time the revenue. We can think of the probability function as a convergence mechanism that guides us to a local optimum and the sequence function as an escape mechanism that makes large jumps and transports us to a different landscape.

6.1.1 Algorithms

All algorithms of the class we are discussing differ only in how often they switch form one function to the other and according to which rule they select nodes to maximize the *FLR*. We propose two algorithms that lie in opposite extremes on how often they apply the sequence lemma.

The first algorithm is a greedy version, where we apply continuously the probability function until we converge to a local maximum for the given sequence. Then based on that probability vector we obtain a new ordering. The algorithm terminates when no improvement is made. The precise description is:

Greedy Local Search

Input: *A graph $G(V,E)$, a probability vector $\mathbf{p}$ and a number ϵ*
Expectation:
1. *Apply the sequence Lemma and obtain a new ordering.*
Maximization:
Repeat
2. *Using a cyclic rule select one variable at a time and update by maximizing the FLR revenue:*

$$p_i = \frac{1}{2} + \frac{1}{2}\frac{\sum_{j>\pi(i)} p_j(1-p_j)w_{ij}}{\sum_{j<\pi(i)} p_j w_{ji} + w_{ii}}$$

until *convergence of $\vec{p}$.*
3. *Repeat **Step 1 and 2** until no improvement in Revenue is made: $\Delta R < \epsilon$.*
Output: *the probability vector $\mathbf{p}$ and the sequence π.*

The previous algorithm is severely biased towards the probability updating rule, i.e. gives more value in obtaining first a better probability vector. On the other extreme we propose an algorithm that alternates between the two functions at each step. The reasoning behind the algorithm is that we are trying to obtain a better sequence simultaneously with updates for the probability vector. In other words we are performing a cautious search making alternating moves in both directions (horizontally and vertically) of the gird.

This algorithm at each step decides which node is to be positioned next in the permutation as well as the probability he should be priced. To accomplish this feature we need two estimates for the probability vector. The probability from the previous iteration(epoch) and a "current" estimate. At each time point we have a closed set, vertices for which we have decided about both their position in the permutation as well as their probability, and an open set which consists of vertices that are candidates for the next position in the permutation. The algorithm proceeds by updating for all candidate nodes their "current" probability estimate by using old values for candidate nodes and the updated values for nodes in the closed set. Our reasoning is that all candidates nodes could be the next node in the permutation and as such we must base our decision calculating the probability that would be assigned to them if they really were considered next. The decision is made on the basis of the conditions that the sequence lemmata provide, i.e. which node has the largest probability from the candidate ones(symmetric case) and which node maximizes the gain between influence lost and gained(asymmetric case). A high level description of this algorithm is:

Cautious Local Search

Input: *A graph $G(V,E)$, a probability vector $\mathbf{p}$ and a number ϵ*
Initialization:
1. *Insert every vertex in a priority queue according to the criterion of the sequence lemma.*
2. *Initialize the Closed Set as $\emptyset$.*
Repeat
3. *u=ExtractMax(Queue).*
4. *Update the probability for u by maximizing the FLR.*
5. *Update the new probability estimates for neighbours of u by maximizing the FLR for each node using the updated probabilities for nodes in the closed set and the initial values for nodes in the queue.*
6. *Based on the new estimates update the values that the sequence lemma requires(keys of nodes).*
until *Queue is Empty.*
If *$\Delta R > \epsilon$* ***go to*** *Step 1.*
Output: *the probability vector $\mathbf{p}$ and the sequence π.*

6.1.2 Convergence

Having thoroughly presented our local search strategies, we discuss the issue of convergence. Observe that both these algorithms are guaranteed to converge to a local optimum since the revenue is bounded from above and our strategies guarantee an increase in revenue at each step. However, reasoning of this kind does not really say anything about how fast we will reach that local optimum. To study convergence rates we must study the update rule:

$$p_i^{(k+1)} = \frac{1}{2} + \frac{1}{2}\left[\frac{\sum_{j>\pi(i)} p_j^{(k)}(1-p_j^{(k)})w_{ij}}{\sum_{j<\pi(i)} p_j^{(k)} w_{ji} + w_{ii}}\right]_{\leq 1} \tag{6.3}$$

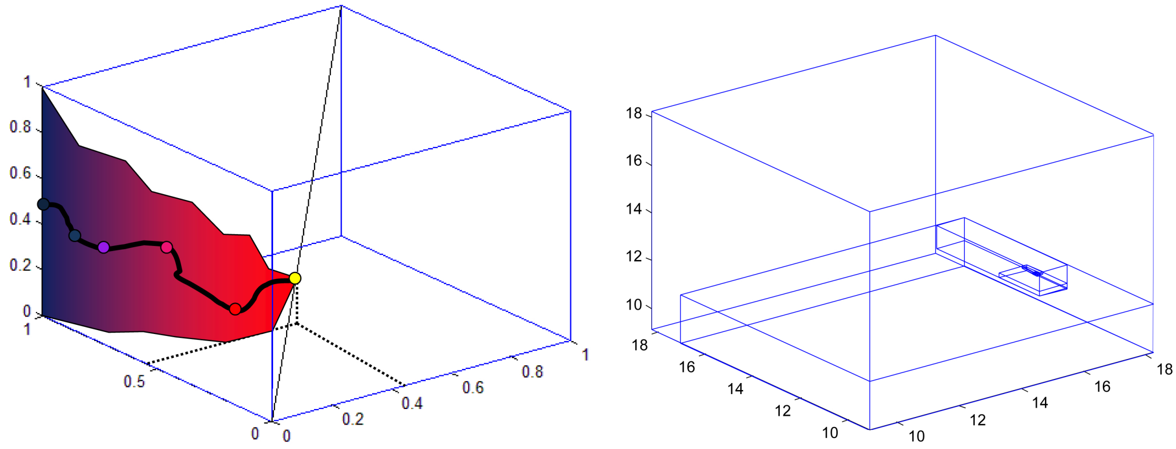


Figure 6.2: (a)Initially, we only know that our variable p_i is in $[0,1]$ and as the algorithm proceeds we get more and more information.(b)Local Search applied to the symmetric model, where the volume of the box indicate the distance to convergence.

It is more convenient to consider the update rule in terms of the discount ratios r_i given by $p_i = 0.5 + 0.5r_i$. The update rule can now be rewritten:

$$r_i^{(k+1)} = f_i(\vec{r}^{(k)}) = \frac{1}{4}\left[\frac{\sum_{j>\pi(i)} w_{ij}(1-(r_j^{(k)})^2)}{\sum_{j<\pi(i)}(0.5+0.5r_j^{(k)})w_{ji} + w_{ii}}\right]_{\leq 1} \tag{6.4}$$

The operator f_i has the nice property of isotonicity $\forall i \in V$. That is if $\vec{x} \geq \vec{y}$ then $f_i(\vec{x}) \leq f_i(\vec{y})$. Futhermore, this kind of update rule is well known in the optimization literature as Cyclic Coordinate Minimization(CCM)[13]. Recently, Saha and Tewari[119] showed that CCM has linear convergence rates when the isotonicity assumption holds[119, Section 4.3, 4.4, 5]. Nevertheless, we provide our own analysis which gives further intuition on why CCM must converge pretty fast.

The idea is that we initiate the discount vector $\vec{r}^{(0)}$ at an unknown location in $[0,1]^n$(full uncertainty) and we virtually iteratively apply the update rule to the unknown vector. Because the vector $\vec{r}^{(0)}$ belongs in $[0,1]^n$ and due to the isotonicity of the update rule f_i, we can obtain more and more refined upper and lower bounds. Specifically, let $\vec{r}^{(k)}$ be the sequence of discount ratios that results after k updates. Our approach is to bound this sequence from above and below respectively by two sequences $\vec{u}^{(k)}$, $\vec{\ell}^{(k)}$ such that $\ell_i^{(k)} \leq r_i^{(k)} \leq u_i^{(k)}$, $\forall\ i \in V$ and for all k, or in matrix notation:

$$\vec{\ell}^{(k)} \leq \vec{r}^{(k)} \leq \vec{u}^{(k)} \tag{6.5}$$

The method can be visualized as applying a clamp for every coordinate(probability) and gradually narrowing the grip until the two ends meet. We initialize our bounds with the obvious choice $\vec{u}^{(0)} = \vec{1}$ and $\vec{\ell}^{(0)} = \vec{0}$. Then we obtain the new bounds by the rules:

$$u_i^{(k+1)} = f_i(\vec{\ell}^{(k)}) \tag{6.6}$$
$$\ell_i^{(k+1)} = f_i(\vec{u}^{(k)}) \tag{6.7}$$

We will show by induction that the sequences $\vec{u}^{(k)}, \vec{\ell}^{(k)}$ converge to the same limit. Observe that if any of the two sequences reach the limit, then the other automatically will reach it as well($\vec{\ell} = f(\vec{\ell}) \Rightarrow u = f(\vec{\ell}) = \vec{\ell} = f(\vec{u})$). We initially assume that $\vec{u}^{(k+1)} \leq \vec{u}^{(k)}$ and $\vec{\ell}^{(k+1)} \geq \vec{\ell}^{(k)}$. If that is the case then we show using the isotonicity property of the update rule that the two sequences are monotone:

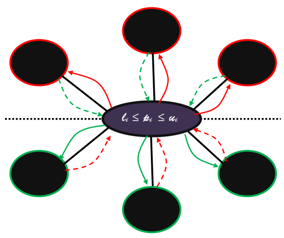
$$u_i^{(k+2)} = f_i(\vec{\ell}^{(k+1)}) \leq f_i(\vec{\ell}^{(k)}) = u_i^{(k+1)} \tag{6.8}$$

$$\ell_i^{(k+2)} = f_i(\vec{u}^{(k+1)}) \geq f_i(\vec{u}^{(k)}) = \ell_i^{(k+1)} \tag{6.9}$$

Now, there are two properties that we must show to complete the proof. Firstly, we must show that the base case holds,i.e. $\vec{u}^{(1)} \leq \vec{u}^{(0)}$ and $\vec{\ell}^{(1)} \geq \vec{\ell}^{(0)}$, and secondly that these inequalities are in fact strict for at least one of the two sequences $\vec{u}, \vec{\ell}$ at each step k.

Our approach in remedying the above situation is to consider the process of cyclic coordinate maximization as a message-passing algorithm. Specifically, each time a node's ratio(probability) is updated it passes "messages" to its neighbours. The concept of a message is that the update of one node will have a contribution on how much the the ratio of an adjacent node will change after we apply the update rule to it. Hence, we can think that a node "accumulates" messages from his neighbours, which are being updated, until it is updated itself. What remains is to quantify the contribution of individual messages.

There are two kinds of messages: *future* messages and *past* messages. *Future* messages are passed from nodes later in the permutation to adjacent nodes that appear earlier in the permutation. The inverse applies for *past* messages.



We analyse first *future messages*. Assume that node i has been just updated and there is a change in value δr_i. However, we have only information for the upper and lower bounds. This means that we should have $\delta u_i < 0$ or $\delta \ell_i > 0$, because if neither holds then that would mean that the truncation option applies for both the lower and upper bound but then that would mean that r_i would have converged. Therefore, if there is a coordinate that hasn't already converged we are guaranteed to have a "message". The change resulting from the message from i to an adjacent node j that is updated is:

$$\ell_j^{(k+1)} - \ell_j^{(k)} = f_j(\vec{u}^{(k)}) - f_j(\vec{u}^{(k-1)}) \geq \frac{1}{4} \frac{w_{ji}}{\sum_{t<\pi(j)} w_{tj} + w_{jj}} \left[(u_i^{(k-1)})^2 - (u_i^{(k)})^2 \right]$$

Therefore, for future messages we can say that:

$$\delta \ell_j^{(k+1)} \geq F_{ij} \delta |u_i^{(k)}| \tag{6.10}$$

$$\delta u_j^{(k+1)} \geq F_{ij} \delta \ell_i^{(k)} \tag{6.11}$$

Respectively, for "past" messages we have:

$$\begin{aligned} \delta \ell_j^{(k+1)} &= \frac{1}{4} \frac{\sum_{t>\pi(j)} w_{jt}(1-(u_t^{(k)})^2)}{\sum_{t<\pi(j)}(0.5+0.5u_t^{(k)})w_{tj}+w_{jj}} - \frac{1}{4} \frac{\sum_{t>\pi(j)} w_{jt}(1-(u_t^{(k-1)})^2)}{\sum_{t<\pi(j)}(0.5+0.5u_t^{(k-1)})w_{tj}+w_{jj}} \\ &\geq \frac{1}{8} \left(\frac{\sum_{t>\pi(j)} w_{jt}(1-(u_t^{(k-1)})^2)}{(\sum_{t<\pi(j)}(0.5+0.5u_t^{(k-1)})w_{tj}+w_{jj})^2} \right) \left[u_i^{(k-1)} - u_i^{(k)} \right] \end{aligned} \tag{6.12}$$

This time however we cant obtain a lower bound for the change that is independent of the sequence of updates. Fortunately, we only need future messages for our purpose and we will ignore past messages as making no difference. Every node receives a number of future messages unless he appears last among his neighbours in the permutation in which case automatically $r_i = u_i = \ell_i = 0$ and thus converges. Hence, what really happens for nodes that have not yet converged is that at each update they sum the messages from "future" neighbours:

$$\delta\ell_j^{(k+1)} \geq \sum_{i>\pi(j)} F_{ij}|\delta u_i^{(k)}| \tag{6.13}$$

$$\delta u_j^{(k+1)} \geq \sum_{i>\pi(j)} F_{ij}\delta\ell_i^{(k)} \tag{6.14}$$

These relation along with the condition that if we have not full convergence there is i such that $\delta\ell_i > 0$ or $|\delta u_i| > 0$, guarantee that at least one of the inequalities (27,28) are in fact strict. This concludes our proof.

6.2 Eigenvector-based Heuristics

In the previous sections we provided approximation algorithm for the problem of Revenue Maximization. We also introduced natural local search strategies to improve upon a given solution. Nevertheless, we were not able to correlate the network position of a buyer with the right price to be offered, as our techniques, except from the personalized version of Influence and Exploit, disregard at large the network structure and are based on random sampling. We propose intelligent heuristics in order to fill that gap.

We have seen that if we have the vector of probabilities then we can find the right sequence and then initiate a local search algorithm to improve our solution. Hence our heuristics are aiming to obtain a "good" estimate of the probability vector. But what is a "good" estimate? We argue that any good estimate should incorporate two characteristics:

- *Closure under Structure:* meaning that it should take into account the complete structure of the graph and the complex recursive dependencies therein.
- *Closure under Optimality:* meaning that in an approximate sense every individual probability estimation of a node should consist a "best response strategy" given the other estimations.

The first condition refers to the combinatorial nature of our problem, i.e. the Max-FAS problem, whereas the second condition corresponds to the pricing aspect. Any reasonable heuristic should take into account this dual nature of our problem.

6.2.1 "Stingy" PageRank

We initially examine the symmetric case and the results obtained here will motivate the techniques introduced for the asymmetric setting. In the beginning of this chapter, we provided conditions that an optimal strategy must meet. Particularly, we required that the probabilities should "satisfy", while remaining in $[0.5, 1]$, the stationary point equations for the expected Revenue:

$$\frac{\partial R}{\partial p_i} = 0 \Leftrightarrow (1-2p_i)\left(\sum_{j\in B_i} w_{ji} + w_{ii}\right) + \left(\sum_{j\in A_i} wijp_j(1-p_j)\right) = 0,\ \forall\, i \in V \tag{6.15}$$

These equations will form the basis of our heuristic. At this point we make some comments that will provide intuition into our approach:

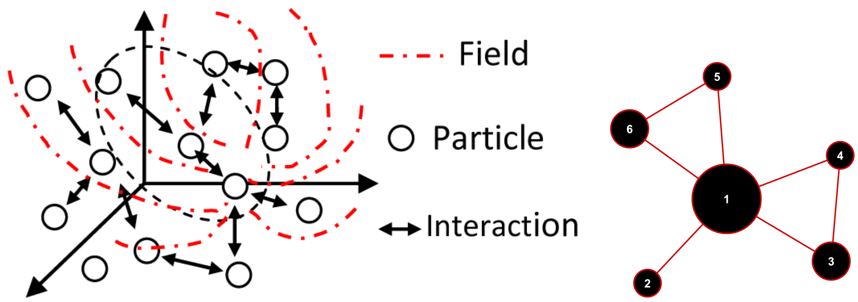


Figure 6.3: (a)Interacting particles conceptualization of Stingy PageRank, where the particles dynamics are advancing the system to equilibrium(closure).(b)We can view the PageRank approach as producing a coarse picture of the role of each node in the Reveneue Maximization setting. The size of the nodes indicates the potential of a node as an influencer.

- *Comment 1:* Given the success of uniform sampling strategies, where the ordering of buyers is random, and the fact that for the Max-FAS problem the best known algorithm is to consider a random permutation, we conclude that a random ordering of vertices is not that bad.
- *Comment 2:* The ratio r_i has the interpretation of a discount offered to node i in order to increase the probability that he accepts and utilize his social position.
- *Comment 3:* Even a fully myopic strategy, where we extract the maximum amount of revenue from each buyer disregarding the network effects, has a constant $1/4$ approximation ratio.

In order to achieve the "Closure" feature under both structure and optimality we use the stationary point equations (6.15) and our estimation for the probability vector will be the output of the following process. Imagine that we have a system of n particles that are affected by a "field" $\vec{f} = (f_1, \ldots, f_n)$, which only depends for each particle from certain adjacent particles(neighbours) and guides our system to equilibrium(closure). In this setting we additionally require that:

1. The *ordering* of particles(buyers) is random(Comment 1).
2. The seller is stingy, that is, he prefers not to give big discounts either because he does not trust the estimations of w_{ij} or in order to discourage foul play. Therefore, he prefers that r_i are kept small(Comment 2).
3. The system is initialized from a fully myopic pricing point: $r_i \simeq 0, \forall\ i \in V$ (Comment 3).

Having provided the motivation to our approach we proceed with the derivation of the heuristic. We note that we consider $w_{ii} = 0$ for simplicity and because this is the most interesting case. However, the techniques used here can be easily extended for the case were $w_{ii} \neq 0$. Because the ordering is random, every vertex has equal probability of being considered before or after any other vertex. Therefore, we consider the *mean field approximation*(in terms of the sequence) of the stationary point equations (6.15):

$$(1 - 2p_i)\left(\frac{1}{2}\sum_{j\neq i} w_{ji}p_j\right) + \frac{1}{2}\sum_{j\neq i} w_{ij}p_j(1 - p_j) = 0 \ \forall i \in V \tag{6.16}$$

If we rewrite (6.16) in terms of the discount ratios and solve for r_i, we get:

$$r_i = f_i(\vec{r}_{-i}) = \frac{1}{4}\frac{\sum_{j\neq i} w_{ij} - \sum_{j\neq i} w_{ij} r_j^2}{\sum_{j\neq i} w_{ji}(0.5 + 0.5 r_j)} \tag{6.17}$$

Finally, we consider "stingy" pricing where the field (6.17) is approximated by a modified field $\vec{f}'$ such that $\vec{f} \geq \vec{f}'$:

$$r_i = f_i(\vec{r}_{-i}) = \frac{1}{4}\frac{\sum_{j\neq i} w_{ij} - \sum_{j\neq i} w_{ij} r_j}{\sum_{j\neq i} w_{ji}} \tag{6.18}$$

where we have used that $r_j \leq 1$ and $(1 - r_j^2) \geq (1 - r_j)$. Before, proceeding with the final result we note that if a vertex has no neighbours we can safely price him with the fully myopic price 1/2 and thus would not be considered in this process, hence we know that $\sum_{j\neq i} w_{ji} \neq 0$ for vertices considered in this process. What we have achieved with (6.18) is essentially a linearisation of the field. This allows us to have a closed self-consistent expression for the "equilibrium point" we are searching for. Finally, we state that *"Stingy" PageRank* is the solution of the system:

$$\vec{r} = -\alpha \vec{W}\vec{D}^{-1} \cdot \vec{r} + \vec{\beta} \tag{6.19}$$

where $\alpha = 1/4$, $\vec{W} = w_{ij}, \vec{D} = \sum_{j\neq i} w_{ji}$ and $\vec{\beta} = 1/4 \cdot \vec{1}$. The name comes from that fact that this is the same formulation that PageRank [?] is computed. We can rewrite (6.19) as linear system:

$$\left(\vec{I} + \alpha \vec{W}\vec{D}^{-1}\right)\vec{r} = \vec{\beta} \tag{6.20}$$

Theorem 13. *"Stingy" PageRank always has a unique solution.*

Proof. Note that the matrix $\vec{W}$ is symmetric with non-negative elements, from SVD theorem we deduce that it has non-negative eigenvalues. Additionally, $\vec{D}^{-1}$ is a diagonal matrix with all elements positive. Therefore, the matrix $(\vec{I} + \alpha\vec{W}\vec{D}^{-1})$ has only positive eigenvalues and the linear system has a unique solution. □

In practice we would apply Stingy PageRank separately for each connected component and normalize the vector $\vec{r}$so that the lower value is 0 and the maximum value is 1. The probability estimation would then be given by $\hat{\vec{p}} = 0.5 + 0.5\vec{r}$. Based on this estimation we could derive the sequence and if needed improve the solution by using one of the local search strategies we proposed.

6.2.2 Hubs and Authorities

The developments for the symmetric setting give us a strong motivation to consider a similar approach for the asymmetric case. In this setting, naturally, we are aiming to utilize the HITS procedure introduced by Kleinberg [81]. Our reasoning is that every node has a dual role, *influencing* other nodes and *being exploited* to extract revenue. This duality is in direct correspondence with Kleinberg's *authorities* and *hubs* concept.

In this setting however we can not hope for achieving complete closure, both structure and optimality, and we will aim only for structural closure. Specifically, we consider for each node two numbers:

1. *Hubness* ratio h_i that expresses how well a node executes his role as a revenue provider.
2. *Authoritativeness* ratio a_i that express how appropriate is a node as an influence.

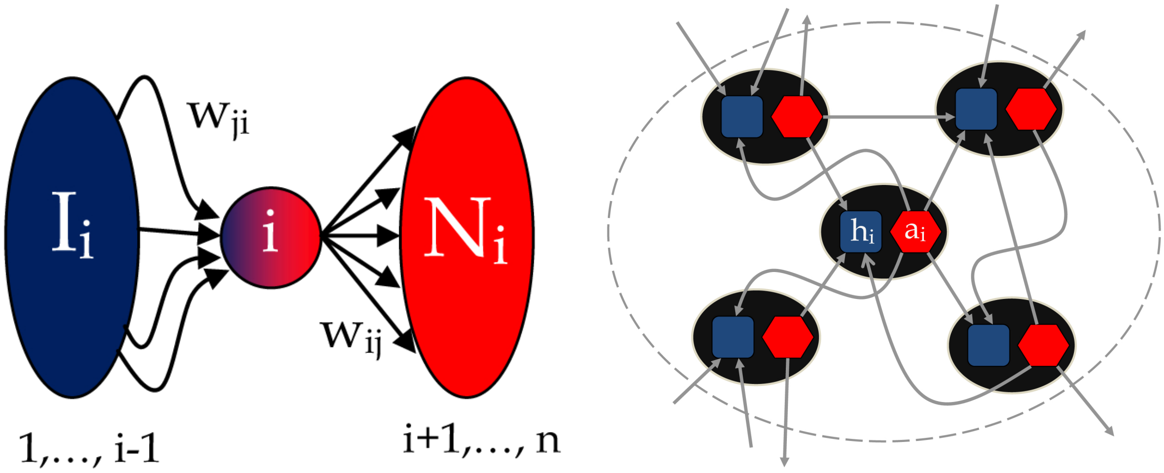


Figure 6.4: (a)Dual value of a node: intrinsic value(blue) and network value(red).(b)Jon Kleinberg's Hubs and Authorities idea.

Based on this concept, we slightly modify the HITS procedure to establish the closure property. Specifically, we consider that we have a random ordering and the ratios must satisfy:

$$a_i = \frac{1}{\sum_{j\neq i} w_{ij}} \sum_{j\neq i} w_{ij} h_j, \ \forall i \in V \tag{6.21}$$

$$h_i = \frac{1}{\sum_{j\neq i} w_{ji}} \sum_{j\neq i} w_{ji} a_j, \ \forall i \in V \tag{6.22}$$

This conditions roughly say that the higher the ratio of a node's influence upon neighbours towards the cumulative influence, the higher is the authoritativeness of the node. Similarly, the higher the authoritativeness of a nodes in-neighbours is, the higher the quality of a node as a revenue provider. These equations can be written in matrix notation as:

$$\vec{a} = \mathbf{D}_{out}^{-1}\vec{W} \cdot \vec{h} \tag{6.23}$$

$$\vec{h} = \mathbf{D}_{in}^{-1}\vec{W}^T \cdot \vec{a} \tag{6.24}$$

These equations actually consist a single eigenvalue problem:

$$\vec{a} = \mathbf{D}_{out}^{-1}\vec{W}\mathbf{D}_{in}^{-1}\vec{W}^T \cdot \vec{a} \tag{6.25}$$

$$\vec{h} = \mathbf{D}_{in}^{-1}\vec{W}^T\mathbf{D}_{out}^{-1}\vec{W} \cdot \vec{h} \tag{6.26}$$

Hence, we could solve either of the two problems and then obtain the other vector using equations (6.23,6.24). Kleinberg showed that the solution to this problem always exisst and consists of the principle eigenvector of the matrix, where all coordinates of the eigenvector are positive due to the Perron-Frobenius Theorem[102].

What remains is to combine the vectors $\vec{a}, \vec{h}$ to a single estimation for the probability vector $\vec{p}$. In the calculation of the authority and hub vectors of Kleinberg there was inherent a normalization procedure. In our case that won't be necessary as we initiate our iterations from a point were $\vec{a} = \vec{h} = \vec{1}$ and in the update rule we have divided with the sum of in-weights and out-weights respectively. Bearing in mind that a vertex with high authority rating should get

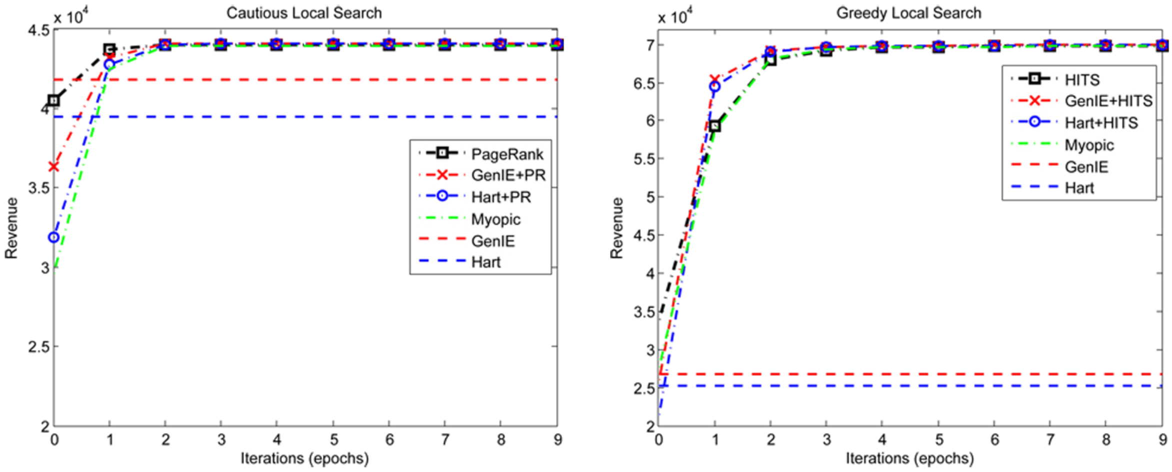


Figure 6.5: Experimental evaluation of eigenvector heuristics. (a)HepPh Collaboration Graph(undirected).(b)Epinions Social Network(directed).

a big discount and a vertex with high hub rating should get a small discount and the fact that these two roles are complementary($a_i + h_i$ should be 1), we consider the following estimation for the probability vector:

$$\hat{p}_i = \frac{1}{2} + \frac{1}{2}\frac{a_i}{a_i + h_i} \tag{6.27}$$

We therefore have provided scalable methods in obtaining an educated guess for the probability vector. This estimation could be used as well in augmenting Influence and Exploit Strategies, that is we could synthesize the two approaches by considering the pricing scheme that the generalized Influence and Exploit strategy requires but instead of a random sequence we could use the sequence obtained by applying the sequence lemmata for the estimation that our heuristics produce.

6.3 Experimental Results

We evaluated our heuristics experimentally. We considered two specific graphs with unity weights: the ArXiV HepPh physics collaboration network($n = 12008$, $m = 237000$)[93] for the undirected case and the Epinions network($n = 75879$, $m = 508837$)[113] for the directed case. We implemented four strategies. The fully myopic one, the IE strategy of Hartline et.al, the Generalized IE and the HITS heuristic. Particularly, in order to avoid tedious sampling issues, we implemented a variation of these IE strategies where the ordering of vertices is according to the HITS heuristic and the pricing follows the distribution that IE strategies require. For instance in the $0.5 - -1$ strategy, $\lceil 0.333n \rceil$ first vertices where given the product for free and $\lfloor 0.667n \rfloor$ where offered the fully myopic price. We calculated the initial revenue of these strategies and then applied a local search algorithm to improve it. We observe that in both cases our heuristics initially outperform the other strategies, while the local search algorithms significantly improve all initial solutions with a small number of iterations. It is interesting that the local search strategies show little sensitivity with respect to the initial solution, as they achieve roughly the same performance for all strategies. We also include with vertical lines the theoretical performance of IE strategies in order to have a benchmark. In all cases local search strategies outperform the theoretical bounds and especially in the directed setting. These results constitute evidence for the efficacy of both our local search strategies and heuristics.

6.4 Pricing via Calculus of Variations

We close this chapter by treating the problem of Revenue Maximization under the framework of Calculus of Variations. Recall the Generalized Influence and Exploit strategies from the previous chapter, where we assign each vertex to a one of the k pricing classes with some probability uniformly for all vertices. If we allow the pricing classes to be defined over the whole interval $[0.5, 1]$ then, instead of looking for appropriate probabilities, we are searching for probability density functions with support $[0.5, 1]$. The optimization problems where the decision variables are functions are the object of Calculus of Variations[51].

6.4.1 Uniform Sampling

We want to find a probability distribution $\mathcal{P} : [0.5, 1] \to [0, 1]$, describing the probability $\mathcal{P}(p_i)$ that any node is assigned the probability of acceptance p_i, that maximizes the expected Revenue:

$$\begin{aligned} \max R(\mathcal{P}) &= \int_{1/2}^{1} x(1-x)\mathcal{P}(x)\left[\frac{1}{2}x\mathcal{P}(x) + \int_{x}^{1} z\mathcal{P}(t)dt\right]dx \\ s.t \quad & \int_{1/2}^{1} \mathcal{P}(x)dx = 1 \\ & \mathcal{P}(x) \geq 0 \end{aligned}$$

Thus we want to optimize a function on a infinite-dimensional space(functions) under convex constraints. This family of problems are handled in the *Calculus of Variations* framework. Specifically, we will transform our problem into a problem of Optimal Control. Let $u(x) = \mathcal{P}(x)$ be the control variable and $y(x) = \int_x^1 t\mathcal{P}(t)dt$, $z(x) = \int_{1/2}^x \mathcal{P}(t)dt$ be the state variables then we get:

$$\begin{aligned} \max \quad & R(u) = \int_{1/2}^{1} \left(x(1-x)u(x)\left[\frac{1}{2}xu(x) + y(x)\right] - u(x)\mu(x) \right) dx \\ s.t \quad & y' = -x \cdot u(x) \\ & z' = u(x) \\ & u(x)\mu(x) = 0, \ \mu(x) \geq 0, \ y(1) = 0, \ z(1/2) = 0, \ z(1) = 1 \end{aligned}$$

Let H be the Hamiltonian of our system:

$$H = x(1-x)u(x)\left[\frac{1}{2}xu(x) + y(x)\right] - u(x)\mu(x) + \lambda_1(x)\left[-xu(x)\right] + \lambda_2(x)u(x)$$

We will solve this problem using the Euler-Lagrange equations:

$$\begin{aligned}
H_u = 0 &\Leftrightarrow x^2(1-x)u + x(1-x)y - \mu - \lambda_1 x + \lambda_2 = 0\\
\lambda_1^{'} = -\frac{\partial H}{\partial y} &\Leftrightarrow \lambda_1^{'} = x(1-x)u\\
\lambda_2^{'} = -\frac{\partial H}{\partial z} &\Leftrightarrow \lambda_2^{'} = 0\\
y^{'} &= -x \cdot u(x)\\
z^{'} &= u(x)\\
u(x)\mu(x) = 0,\ \mu(x) \geq 0 &\Rightarrow u(x) > 0 \Rightarrow \mu(x) = 0
\end{aligned}$$

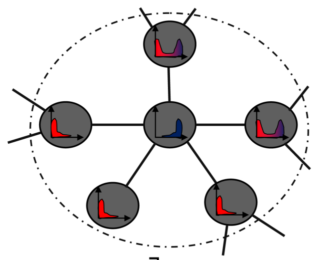
with the necessary initial and terminal conditions:

$$\lambda_1(1/2) = 0,\ y(1) = 0,\ z(1/2) = 0,\ z(1) = 1$$

We have ended up with a Two Point Boundary Value Problem (TPBVP). We may not be able to solve it analytically but there are well developed numerical methods for solving such problems, see *Applied Optimal Control*[24].

6.4.2 Eigen-Distributions

The above discussion motivates us to generalize our approach and instead of considering a single probability distribution common for every node i, we can consider that every node i is assigned a probability of acceptance x according to his own distribution $\mathcal{P}_i(x)$. Let N_i be the set of indices that are adjacent to node i ,then the expected Revenue that we exert from node i is:



$$R_i = \int_{1/2}^{1} x(1-x)\mathcal{P}_i(x)\left[\frac{1}{2}x\sum_{j\in N_i} A_{ji}\mathcal{P}_j(x) + \sum_{j\in N_i} A_{ji}\int_x^1 tP_j(t)dt\right]dx$$

Now the total revenue can be expressed as a sum of the expected revenue from each node $R = \sum_{i\in V} R_i$. Let S be the space of a valid probability distribution with support on the domain $[1/2, 1]$, we want to optimize the total expected revenue on the product space S^n:

$$\begin{aligned}
\max\ R(\mathcal{P}_1,\ldots,\mathcal{P}_n) &= \int_{1/2}^{1}\sum_{i\in V} x(1-x)\mathcal{P}_i(x)\left[\frac{1}{2}x\sum_{j\in N_i} A_{ji}\mathcal{P}_j(x) + \sum_{j\in N_i} A_{ji}\int_x^1 t\mathcal{P}_j(t)dt\right]dx\\
s.t \quad & \int_{1/2}^{1}\mathcal{P}_i(x)dx = 1,\ \mathcal{P}_i(x) \geq 0, \forall i \in V.
\end{aligned}$$

Extending our treatment of the single distribution setting we define $y_i(x) = \int_x^1 t\mathcal{P}_j(t)dt$, $z_i(x) = \int_{1/2}^{x}\mathcal{P}_j(t)dt$ and the control variables $u_i(x) = \mathcal{P}_i(x)$. We calculate the Hamiltonian anew:

$$H = x(1-x)\sum_{i\in V} u_i(x)\left(\frac{1}{2}\sum_{j\in N_i} A_{ji}u_j + \sum_{j\in N_i} A_{ji}y_j\right) + \sum_{i\in V}\lambda_{i,2}u_i - \lambda_{i,1}xu_i(x) - \mu_i(x)u_i(x)$$

Applying the Euler-Lagrange equations for every $i \in V$ we get:

$$\begin{aligned}
x^2(1-x)\sum_{j\in N_i} A_{ji}u_j &+ x(1-x)\sum_{j\in N_i} A_{ji}y_j - \mu_i - \lambda_{i,1}x + \lambda_{i,2} = 0 \\
\lambda_1^{'} &= -x(1-x)\sum_{j\in N_i} A_{ij}u_j, \\
\lambda_{i,2}^{'} &= 0 \\
y_i^{'} &= -x\cdot u_i(x) \\
z_i^{'} &= u_i(x) \\
u_i(x)\mu_i(x) = 0,\ \mu_i(x) \geq 0 &\Rightarrow u_i(x) > 0 \Rightarrow \mu_i(x) = 0
\end{aligned}$$

with the necessary initial and terminal conditions:

$$\lambda_{i,1}(1/2) = 0,\ y_i(1) = 0,\ z_i(1/2) = 0,\ z_i(1) = 1,\ \forall i \in V$$

In this setting we have embedded the actual graph structure in the above dynamical system. We can solve the problem numerically only if all the eigenvalues of the adjacency matrix are non-zero, as we can calculate the u_i's only from the system of linear equations that arises from the first set equations ($\partial H/\partial u_i = 0$). Actually, it suffices for the eigenvalues of the corresponding adjacency matrices of the connected components to have positive eigenvalues.

6.4.3 Eigen-Generalized Influence and Exploit

A natural alternative to working in a infinite dimensional space is to consider a finite discretization $p_i = \mathcal{P}(x_i)$ of the distribution $\mathcal{P}(x)$ where x_i could be for instance $x_i = 1 - \frac{i-1}{2*(m-1)}$ and m the number of "buckets". Now the problem can be written:

Motivated by the first section and to obtain a graph specific pricing algorithm, we consider a discretized probability distribution p_k^i for each node i. Specifically, the problem is formulated as:

$$\begin{aligned}
\max\ R(\mathbf{p}^1,\ldots,\mathbf{p}^n) &= \sum_{i=1}^{n}\sum_{k=1}^{m} x_k(1-x_k)p_k^i\left(\frac{1}{2}x_k\sum_{j\neq i} A_{ji}p_k^j + \sum_{r=1}^{k-1} x_r \sum_{j\neq i} A_{ji}p_r^j\right) \\
s.t \quad & \sum_{k=1}^{m} p_k^i = 1,\ p_k^i \geq 0,\ k = 1,\ldots,m,\ i = 1,\ldots,n
\end{aligned}$$

Again we have the same Quadratic structure for the objective function. Specifically the matrix $Q = \left[\frac{\partial^2 R}{\partial p_k^i \partial p_\ell^j}\right]$:

$$\begin{aligned}
\frac{\partial^2 R}{\partial p_k^i \partial p_\ell^j} &= x_k(1-x_k)x_\ell A_{ji},\ \ell < k \\
\frac{\partial^2 R}{\partial p_k^i \partial p_k^j} &= x_k(1-x_k)x_k A_{ji},\ \ell = k \\
\frac{\partial^2 R}{\partial p_k^i \partial p_\ell^j} &= x_\ell(1-x_\ell)x_k A_{ij},\ \ell > k
\end{aligned}$$

Thus, we could initialize all the distributions with the solution we derived from the uniform case which would guarantee us a 0.7059 approximation ratio and therefore from the solution of the Quadratic program obtain an improved *node specific* solution. We have developed a quite general framework for designing pricing strategies both in the continuous setting as well as in the discrete setting.

Chapter 7

Conclusion

In this thesis, we studied the Revenue Maximization problem under a very specific model of marketing and positive externalities. The main assumptions made are:(i) buyers are approached individually in a sequence by the seller, (ii) the seller implements discriminative pricing, (iii) the influence is additive. The first assumption although leading to marketing strategies that may be implementable(through promotional emails or messages), is in reality impractical as buyers are in general reluctant to respond positively to such offers and even if they do so there would be a significant delay, rendering the sequential promotion process too slow. Discriminative pricing, on the other hand, is known to lead to negative reactions from buyers especially when it happens in their close social circle. Finally, the additive influence assumption is a very rough approximation to buyers valuations which are known to be governed by a diminishing returns property(concave or submodular functions).

Considering the first two issues, we propose a way by which they can be mitigated even in this model of sequential discriminative marketing. Sequential marketing suffers from the inherent problem that the time to execute a marketing strategy is proportional to the number of buyers n, which usually is very large. A reasonable constraint on the steps executed by a marketing strategy could be that they should be at most a logarithmic function($O(\log n)$) of the number of buyers. This constraint naturally imposes a marketing strategy that at each step the number of buyers considered is geometrically increasing, i.e. at each step we consider constant times(e.g. twice as many) more buyers than the previous. Therefore, we cluster buyers in $O(log(n))$ groups and make simultaneous offers to buyers in each group. The clustering can be made either by adapting appropriately random sampling(Generalized IE), local search procedures[110, 50] or by ranking through Eigenvector Centrality("Stingy PageRank") techniques. This method might reduce the execution time of the marketing strategy but the negative effects of discriminative pricing remain at large.

To alleviate the negative effects of discriminative pricing and still be able to preserve some control over the marketing process, the Influence-and-Exploit idea is apposite. Offering the product for free to some buyers and offering a regular price to the rest is an accepted by consumers marketing practice. Furthermore, it is easily implementable via a gift and posted price mechanism(same price available to all buyers). The Influence-and-Exploit strategy is a clever way to find a good solution to complex problems by breaking the symmetry between the decision variables in order to exploit partial knowledge and intuition about the structure of the optimal solution. Some instances where this idea has been succesfully applied are [27, 65].

The Revenue Maximization and Influence Maximization problems study the same process from different viewpoints, that of value and information propagation respectively. However, in reality these two aspects are intertwined and it is an interesting direction to merge them under one model. Towards, that direction the Linear Threshold Model looks quite promising as it could incorporate the dynamics by considering for each vertex two thresholds θ_V, θ_I (about value and information respectively) and a joint distribution over them F_{P_t,S_t} depending on the price trajectory $P_t = (p_0, \ldots, p_t)$ and the adoption trajectory $S_t = (s_0, \ldots, s_t)$ of its neighbours, where p_t is the posted price and s_t is the set of neighbours that own the product at time t. The goal would be to exploit the knowledge of the dependencies between buyer to design a marketing strategy that maximizes the expected profit. The marketing strategy would be comprised by three parts. A mass marketing campaign with cost which is a non-decreasing function of the probability that a random buyer is informed about the product. A promotion process where a set of buyers are given the product for free with possibly variable cost for each buyer and a posted price trajectory according to which an arbitrary buyer can buy the product at each time point. If these challenges are(or can be) met, then we would obtain a truly Algorithmic Theory of Marketing.

List of Figures

Bibliography

[1] Gagan Aggarwal, Jon Feldman, S. Muthukrishnan, and Martin Pál. Sponsored search auctions with markovian users. In *WINE*, pages 621–628, 2008.

[2] Hessameddin Akhlaghpour, Mohammad Ghodsi, Nima Haghpanah, Vahab S. Mirrokni, Hamid Mahini, and Afshin Nikzad. Optimal iterative pricing over social networks. In *Proceedings of the 6th international conference on Internet and network economics*, WINE'10, pages 415–423, Berlin, Heidelberg, 2010. Springer-Verlag.

[3] Farid Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization*, 5:13–51, 1993.

[4] Noga Alon, Michal Feldman, Ariel D. Procaccia, and Moshe Tennenholtz. A note on competitive diffusion through social networks. *Inf. Process. Lett.*, 110(6):221–225, February 2010.

[5] Nima Anari, Shayan Ehsani, Mohammad Ghodsi, Nima Haghpanah, Nicole Immorlica, Hamid Mahini, and Vahab S. Mirrokni. Equilibrium pricing with positive externalities. In *Proceedings of the 6th international conference on Internet and network economics*, WINE'10, pages 424–431, Berlin, Heidelberg, 2010. Springer-Verlag.

[6] Krzysztof R. Apt and Evangelos Markakis. Diffusion in social networks with competing products. In *Proceedings of the 4th international conference on Algorithmic game theory*, SAGT'11, pages 212–223, Berlin, Heidelberg, 2011. Springer-Verlag.

[7] Aristotle. *The Nicomachean Ethics (Oxford World's Classics)*. Oxford Paperbacks, April 1998.

[8] David Arthur, Rajeev Motwani, Aneesh Sharma, and Ying Xu. Pricing strategies for viral marketing on social networks. In *Proceedings of the 5th International Workshop on Internet and Network Economics*, WINE '09, pages 101–112, Berlin, Heidelberg, 2009. Springer-Verlag.

[9] Lars Backstrom, Daniel P. Huttenlocher, Jon M. Kleinberg, and Xiangyang Lan. Group formation in large social networks: membership, growth, and evolution. In *KDD*, pages 44–54, 2006.

[10] Maria-Florina Balcan and Nicholas J.A. Harvey. Learning submodular functions. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC '11, pages 793–802, New York, NY, USA, 2011. ACM.

[11] Frank M. Bass. A new product growth for model consumer durables. *Manage. Sci.*, 50(12 Supplement):1825–1832, December 2004.

[12] Bonnie Berger and Peter W. Shor. Approximation alogorithms for the maximum acyclic subgraph problem. In *Proceedings of the first annual ACM-SIAM symposium on Discrete algorithms*, SODA '90, pages 236–243, Philadelphia, PA, USA, 1990. Society for Industrial and Applied Mathematics.

[13] Dimitri P. Bertsekas and Dimitri P. Bertsekas. *Nonlinear Programming.* Athena Scientific, 2nd edition, September 1999.

[14] Shishir Bharathi, David Kempe, and Mahyar Salek. Competitive influence maximization in social networks. In *Proceedings of the 3rd international conference on Internet and network economics*, WINE'07, pages 306–311, Berlin, Heidelberg, 2007. Springer-Verlag.

[15] Sayan Bhattacharya, Janardhan Kulkarni, Kamesh Munagala, and Xiaoming Xu. On allocations with negative externalities. In *WINE*, pages 25–36, 2011.

[16] Philippe Biane, Richard Durrett, and Rick Durrett. Ten lectures on particle systems. In *Lectures on Probability Theory*, volume 1608 of *Lecture Notes in Mathematics*, pages 97–201. Springer Berlin / Heidelberg, 1995. 10.1007/BFb0095747.

[17] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics).* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[18] Liad Blumrosen and Shahar Dobzinski. Welfare maximization in congestion games. In *Proceedings of the 7th ACM conference on Electronic commerce*, EC '06, pages 52–61, New York, NY, USA, 2006. ACM.

[19] Béla Bollobás and Oliver Riordan. *Percolation.* Cambridge University Press, 2006.

[20] Phillip Bonacich. Power and Centrality: A Family of Measures. *American Journal of Sociology*, 92(5):1170–1182, 1987.

[21] Francesco Bonchi, Carlos Castillo, Aristides Gionis, and Alejandro Jaimes. Social network analysis and mining for business applications. *ACM Trans. Intell. Syst. Technol.*, 2(3):22:1–22:37, May 2011.

[22] Allan Borodin, Yuval Filmus, and Joel Oren. Threshold models for competitive influence in social networks. In *Proceedings of the 6th international conference on Internet and network economics*, WINE'10, pages 539–550, Berlin, Heidelberg, 2010. Springer-Verlag.

[23] Felix Brandt, Tuomas Sandholm, and Yoav Shoham. Spiteful bidding in sealed-bid auctions. In *Proceedings of the 20th international joint conference on Artifical intelligence*, IJCAI'07, pages 1207–1214, San Francisco, CA, USA, 2007. Morgan Kaufmann Publishers Inc.

[24] A. E. Bryson and Y. C. Ho. *Applied Optimal Control.* Blaisdell, New York, 1969.

[25] Ozan Candogan, Kostas Bimpikis, and Asuman Ozdaglar. Optimal pricing in the presence of local network effects. In *Proceedings of the 6th international conference on Internet and network economics*, WINE'10, pages 118–132, Berlin, Heidelberg, 2010. Springer-Verlag.

[26] Tim Carnes, Chandrashekhar Nagarajan, Stefan M. Wild, and Anke van Zuylen. Maximizing influence in a competitive social network: a follower's perspective. In *Proceedings of the ninth international conference on Electronic commerce*, ICEC '07, pages 351–360, New York, NY, USA, 2007. ACM.

[27] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. On the advantage over random for maximum acyclic subgraph. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '07, pages 625–633, Washington, DC, USA, 2007. IEEE Computer Society.

[28] Wei Chen, Alex Collins, Rachel Cummings, Te Ke, Zhenming Liu, David Rincón, Xiaorui Sun, Yajun Wang, Wei Wei, and Yifei Yuan. Influence maximization in social networks when negative opinions may emerge and propagate. In *SDM*, pages 379–390, 2011.

[29] Wei Chen, Chi Wang, and Yajun Wang. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '10, pages 1029–1038, New York, NY, USA, 2010. ACM.

[30] Wei Chen, Yajun Wang, and Siyu Yang. Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '09, pages 199–208, New York, NY, USA, 2009. ACM.

[31] Wei Chen, Yifei Yuan, and Li Zhang. Scalable influence maximization in social networks under the linear threshold model. In *Proceedings of the 2010 IEEE International Conference on Data Mining*, ICDM '10, pages 88–97, Washington, DC, USA, 2010. IEEE Computer Society.

[32] Flavio Chierichetti, Jon M. Kleinberg, and David Liben-Nowell. Reconstructing patterns of information diffusion from incomplete observations. In *NIPS*, pages 792–800, 2011.

[33] Vincent Conitzer and Tuomas Sandholm. Expressive markets for donating to charities. *Artif. Intell.*, 175(7-8):1251–1271, 2011.

[34] Vincent Conitzer and Tuomas Sandholm. Computing optimal outcomes under an expressive representation of settings with externalities. *J. Comput. Syst. Sci.*, 78(1):2–14, 2012.

[35] Florin Constantin, Malvika Rao, Chien-Chung Huang, and David C. Parkes. On expressing value externalities in position auctions. In *AAAI*, 2011.

[36] Dan Cosley, Daniel P. Huttenlocher, Jon M. Kleinberg, Xiangyang Lan, and Siddharth Suri. Sequential influence models in social networks. In *ICWSM*, 2010.

[37] Peter Cramton, Yoav Shoham, and Richard Steinberg, editors. *Combinatorial Auctions*. MIT Press, 2006.

[38] David J. Crandall, Dan Cosley, Daniel P. Huttenlocher, Jon M. Kleinberg, and Siddharth Suri. Feedback effects between similarity and social influence in online communities. In *KDD*, pages 160–168, 2008.

[39] Easley David and Kleinberg Jon. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, New York, NY, USA, 2010.

[40] Changrong Deng and Sasa Pekec. Money for nothing: exploiting negative externalities. In *Proceedings of the 12th ACM conference on Electronic commerce*, EC '11, pages 361–370, New York, NY, USA, 2011. ACM.

[41] Xiaotie Deng and Jiajin Yu. A new ranking scheme of the gsp mechanism with markovian users. In *Proceedings of the 5th International Workshop on Internet and Network Economics*, WINE '09, pages 583–590, Berlin, Heidelberg, 2009. Springer-Verlag.

[42] Nikhil R. Devanur and Sham M. Kakade. The price of truthfulness for pay-per-click auctions. In *Proceedings of the 10th ACM conference on Electronic commerce*, EC '09, pages 99–106, New York, NY, USA, 2009. ACM.

[43] Jared M. Diamond. *Guns, Germs, and Steel: The Fates of Human Societies*. W.W. Norton & Co., 1 edition, April 1999.

[44] Pedro Domingos and Matt Richardson. Mining the network value of customers. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '01, pages 57–66, New York, NY, USA, 2001. ACM.

[45] Uriel Feige and Michel X. Goemans. Aproximating the value of two prover proof systems, with applications to max 2sat and max dicut. In *ISTCS*, pages 182–189, 1995.

[46] Uriel Feige, Vahab S. Mirrokni, and Jan Vondrak. Maximizing non-monotone submodular functions. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '07, pages 461–471, Washington, DC, USA, 2007. IEEE Computer Society.

[47] Dimitris Fotakis, Vasilis Gkatzelis, Alexis C. Kaporis, and Paul G. Spirakis. The impact of social ignorance on weighted congestion games. In *Proceedings of the 5th International Workshop on Internet and Network Economics*, WINE '09, pages 316–327, Berlin, Heidelberg, 2009. Springer-Verlag.

[48] Dimitris Fotakis, Piotr Krysta, and Orestis Telelis. Externalities among advertisers in sponsored search. In *Proceedings of the 4th international conference on Algorithmic game theory*, SAGT'11, pages 105–116, Berlin, Heidelberg, 2011. Springer-Verlag.

[49] András Frank. Submodular functions in graph theory. *Discrete Math.*, 111(1-3):231–243, February 1993.

[50] Daya Ram Gaur, Ramesh Krishnamurti, and Rajeev Kohli. The capacitated max k-cut problem. *Math. Program.*, 115(1):65–72, May 2008.

[51] I. M. Gelfand and S. V. Fomin. *Calculus of Variations (Dover Books on Mathematics)*. Dover Publications, October 2000.

[52] Shayan Oveis Gharan and Jan Vondrák. Submodular maximization by simulated annealing. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '11, pages 1098–1116. SIAM, 2011.

[53] Ioannis Giotis and Anna R. Karlin. On the equilibria and efficiency of the gsp mechanism in keyword auctions with externalities. In *Proceedings of the 4th International Workshop on Internet and Network Economics*, WINE '08, pages 629–638, Berlin, Heidelberg, 2008. Springer-Verlag.

[54] Michel X. Goemans, Nicholas J. A. Harvey, Satoru Iwata, and Vahab Mirrokni. Approximating submodular functions everywhere. In *Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '09, pages 535–544, Philadelphia, PA, USA, 2009. Society for Industrial and Applied Mathematics.

[55] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, November 1995.

[56] J. Goldenberg, B. Libai, and E. Muller. Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth. *Marketing Letters*, pages 211–223, August 2001.

[57] Renato Gomes, Nicole Immorlica, and Evangelos Markakis. Externalities in keyword auctions: An empirical and theoretical assessment. In *Proceedings of the 5th International Workshop on Internet and Network Economics*, WINE '09, pages 172–183, Berlin, Heidelberg, 2009. Springer-Verlag.

[58] Manuel Gomez-Rodriguez, Jure Leskovec, and Andreas Krause. Inferring networks of diffusion and influence. *ACM Trans. Knowl. Discov. Data*, 5(4):21:1–21:37, February 2012.

[59] Amit Goyal, Francesco Bonchi, and Laks V.S. Lakshmanan. Learning influence probabilities in social networks. In *Proceedings of the third ACM international conference on Web search and data mining*, WSDM '10, pages 241–250, New York, NY, USA, 2010. ACM.

[60] Amit Goyal, Wei Lu, and Laks V. S. Lakshmanan. Simpath: An efficient algorithm for influence maximization under the linear threshold model. In *ICDM*, pages 211–220, 2011.

[61] Amit Goyal, Wei Lu, and Laks V.S. Lakshmanan. Celf++: optimizing the greedy algorithm for influence maximization in social networks. In *Proceedings of the 20th international conference companion on World wide web*, WWW '11, pages 47–48, New York, NY, USA, 2011. ACM.

[62] Mark Granovetter. Threshold Models of Collective Behavior. *American Journal of Sociology*, 83(6):1420–1443, 1978.

[63] Mark Granovetter. The impact of social structure on economic outcomes. *The Journal of Economic Perspectives*, 19(1):pp. 33–50, 2005.

[64] Venkatesan Guruswami, Rajsekar Manokaran, and Prasad Raghavendra. Beating the random ordering is hard: Inapproximability of maximum acyclic subgraph. In *Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '08, pages 573–582, Washington, DC, USA, 2008. IEEE Computer Society.

[65] Nima Haghpanah, Nicole Immorlica, Vahab Mirrokni, and Kamesh Munagala. Optimal auctions with positive network externalities. In *Proceedings of the 12th ACM conference on Electronic commerce*, EC '11, pages 11–20, New York, NY, USA, 2011. ACM.

[66] Jason Hartline, Vahab Mirrokni, and Mukund Sundararajan. Optimal marketing strategies over social networks. In *Proceedings of the 17th international conference on World Wide Web*, WWW '08, pages 189–198, New York, NY, USA, 2008. ACM.

[67] Refael Hassin and Shlomi Rubinstein. Approximations for the maximum acyclic subgraph problem. *Information Processing Letters*, 51:133–140, 1994.

[68] Wassily Hoeffding. Probability Inequalities for Sums of Bounded Random Variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[69] Peter D. Hoff. Multiplicative latent factor models for description and prediction of social networks. *Comput. Math. Organ. Theory*, 15(4):261–272, December 2009.

[70] Nicole Immorlica, Jon Kleinberg, Mohammad Mahdian, and Tom Wexler. The role of compatibility in the diffusion of technologies through social networks. In *Proceedings of the 8th ACM conference on Electronic commerce*, EC '07, pages 75–83, New York, NY, USA, 2007. ACM.

[71] Satoru Iwata. Submodular function minimization. *Math. Program.*, 112(1):45–64, July 2007.

[72] Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, Filip Radlinski, and Geri Gay. Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. *ACM Trans. Inf. Syst.*, 25(2), April 2007.

[73] Kyomin Jung, Wooram Heo, and Wei Chen. Irie: A scalable influence maximization algorithm for independent cascade model and its extensions. *CoRR*, abs/1111.4795, 2011.

[74] Michael L. Katz and Carl Shapiro. Technology Adoption in the Presence of Network Externalities. *The Journal of Political Economy*, 94(4):822–841, 1986.

[75] Michael L Katz and Carl Shapiro. Product introduction with network externalities. *Journal of Industrial Economics*, 40(1):55–83, March 1992.

[76] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '03, pages 137–146, New York, NY, USA, 2003. ACM.

[77] David Kempe and Mohammad Mahdian. A cascade model for externalities in sponsored search. In *Proceedings of the 4th International Workshop on Internet and Network Economics*, WINE '08, pages 585–596, Berlin, Heidelberg, 2008. Springer-Verlag.

[78] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, STOC '02, pages 767–775, New York, NY, USA, 2002. ACM.

[79] Myunghwan Kim and Jure Leskovec. Multiplicative attribute graph model of real-world networks. *Internet Mathematics*, 8(1-2):113–160, 2012.

[80] Masahiro Kimura and Kazumi Saito. Tractable models for information diffusion in social networks. In *Proceedings of the 10th European conference on Principle and Practice of Knowledge Discovery in Databases*, PKDD'06, pages 259–271, Berlin, Heidelberg, 2006. Springer-Verlag.

[81] Jon M. Kleinberg. Authoritative sources in a hyperlinked environment. *J. ACM*, 46(5):604–632, September 1999.

[82] Jon M. Kleinberg. Information networks, link analysis, and temporal dynamics. In *IFIP TCS*, pages 23–25, 2002.

[83] Dirk Koschützki, Katharina Anna Lehmann, Leon Peeters, Stefan Richter, Dagmar Tenfelde-Podehl, and Oliver Zlotowski. Centrality indices. In *Network Analysis*, pages 16–61, 2004.

[84] Gueorgi Kossinets, Jon M. Kleinberg, and Duncan J. Watts. The structure of information pathways in a social communication network. In *KDD*, pages 435–443, 2008.

[85] Elias Koutsoupias and Christos H. Papadimitriou. Worst-case equilibria. In *STACS*, pages 404–413, 1999.

[86] Piotr Krysta, Tomasz Michalak, Tuomas Sandholm, and Michael Wooldridge. Combinatorial auctions with externalities. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1 - Volume 1*, AAMAS '10, pages 1471–1472, Richland, SC, 2010. International Foundation for Autonomous Agents and Multiagent Systems.

[87] Ariel Kulik, Hadas Shachnai, and Tami Tamir. Maximizing submodular set functions subject to multiple linear constraints. In *Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '09, pages 545–554, Philadelphia, PA, USA, 2009. Society for Industrial and Applied Mathematics.

[88] Jon Lee, Vahab S. Mirrokni, Viswanath Nagarajan, and Maxim Sviridenko. Maximizing nonmonotone submodular functions under matroid or knapsack constraints. *SIAM J. Discret. Math.*, 23(4):2053–2078, January 2010.

[89] Jure Leskovec. Social media analytics: tracking, modeling and predicting the flow of information through networks. In *WWW (Companion Volume)*, pages 277–278, 2011.

[90] Jure Leskovec, Lada A. Adamic, and Bernardo A. Huberman. The dynamics of viral marketing. *ACM Trans. Web*, 1(1), May 2007.

[91] Jure Leskovec, Lars Backstrom, and Jon M. Kleinberg. Meme-tracking and the dynamics of the news cycle. In *KDD*, pages 497–506, 2009.

[92] Jure Leskovec, Daniel P. Huttenlocher, and Jon M. Kleinberg. Predicting positive and negative links in online social networks. In *WWW*, pages 641–650, 2010.

[93] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Trans. Knowl. Discov. Data*, 1(1), March 2007.

[94] Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, and Natalie Glance. Cost-effective outbreak detection in networks. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '07, pages 420–429, New York, NY, USA, 2007. ACM.

[95] Jure Leskovec, Mary McGlohon, Christos Faloutsos, Natalie S. Glance, and Matthew Hurst. Patterns of cascading behavior in large blog graphs. In *SDM*, 2007.

[96] Jure Leskovec, Ajit Singh, and Jon M. Kleinberg. Patterns of influence in a recommendation network. In *PAKDD*, pages 380–389, 2006.

[97] Sai-Ming Li, Mohammad Mahdian, and R. Preston McAfee. Value of learning in sponsored search auctions. In *Proceedings of the 6th international conference on Internet and network economics*, WINE'10, pages 294–305, Berlin, Heidelberg, 2010. Springer-Verlag.

[98] David Liben-Nowell and Jon M. Kleinberg. The link-prediction problem for social networks. *JASIST*, 58(7):1019–1031, 2007.

[99] Thomas M. Liggett. *Interacting Particle Systems*. Springer Berlin Heidelberg, 1985.

[100] Jingfeng Lu. Optimal auctions with asymmetric financial externalities. *Games and Economic Behavior*, 74(2):561 – 575, 2012.

[101] R.K. Merton and N.W. Storer. *The Sociology of Science: theoretical and empirical investigations*. University of Chicago Press, 1973.

[102] Carl D. Meyer, editor. *Matrix analysis and applied linear algebra*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.

[103] Mayur Mohite and Y. Narahari. Incentive compatible influence maximization in social networks and application to viral marketing. In *The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 3*, AAMAS '11, pages 1081–1082, Richland, SC, 2011. International Foundation for Autonomous Agents and Multiagent Systems.

[104] Stephen Morris. Contagion. *The Review of Economic Studies*, 67:57–78, 2000.

[105] Elchanan Mossel and Sebastien Roch. On the submodularity of influence in social networks. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, STOC '07, pages 128–134, New York, NY, USA, 2007. ACM.

[106] Seth A. Myers and Jure Leskovec. On the convexity of latent social network inference. In *NIPS*, pages 1741–1749, 2010.

[107] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher. An analysis of approximations for maximizing submodular set functions—I. *Mathematical Programming*, 14(1):265–294, December 1978.

[108] Alantha Newman. The maximum acyclic subgraph problem and degree-3 graphs. In *Proceedings of the 4th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems and 5th International Workshop on Randomization and Approximation Techniques in Computer Science: Approximation, Randomization and Combinatorial Optimization*, APPROX '01/RANDOM '01, pages 147–158, London, UK, UK, 2001. Springer-Verlag.

[109] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, New York, NY, USA, 2007.

[110] James B. Orlin, Abraham P. Punnen, and Andreas S. Schulz. Approximate local search in combinatorial optimization. *SIAM J. Comput.*, 33(5):1201–1214, May 2004.

[111] C. H. Papadimitriou, A. A. Schäffer, and M. Yannakakis. On the complexity of local search. In *Proceedings of the twenty-second annual ACM symposium on Theory of computing*, STOC '90, pages 438–445, New York, NY, USA, 1990. ACM.

[112] Nishith Pathak, Arindam Banerjee, and Jaideep Srivastava. A generalized linear threshold model for multiple cascades. In *Proceedings of the 2010 IEEE International Conference on Data Mining*, ICDM '10, pages 965–970, Washington, DC, USA, 2010. IEEE Computer Society.

[113] Matthew Richardson, Rakesh Agrawal, and Pedro Domingos. Trust management for the semantic web. In *International Semantic Web Conference*, pages 351–368, 2003.

[114] Matthew Richardson and Pedro Domingos. Mining knowledge-sharing sites for viral marketing. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '02, pages 61–70, New York, NY, USA, 2002. ACM.

[115] E. Rogers. *Diffusion of innovations*. Free Press., New York,, 1995.

[116] Daniel M. Romero, Brendan Meeder, and Jon M. Kleinberg. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In *WWW*, pages 695–704, 2011.

[117] Eldar Sadikov, Montserrat Medina, Jure Leskovec, and Hector Garcia-Molina. Correcting for missing data in information cascades. In *Proceedings of the fourth ACM international conference on Web search and data mining*, WSDM '11, pages 55–64, New York, NY, USA, 2011. ACM.

[118] Eldar Sadikov, Montserrat Medina, Jure Leskovec, and Hector Garcia-Molina. Correcting for missing data in information cascades. In *WSDM*, pages 55–64, 2011.

[119] Ankan Saha and Ambuj Tewari. On the finite time convergence of cyclic coordinate descent methods. *CoRR*, abs/1005.2146, 2010.

[120] Kazumi Saito, Masahiro Kimura, Kouzou Ohara, and Hiroshi Motoda. Learning continuous-time information diffusion model for social behavioral data analysis. In *ACML*, pages 322–337, 2009.

[121] Kazumi Saito, Ryohei Nakano, and Masahiro Kimura. Prediction of information diffusion probabilities for independent cascade model. In *Proceedings of the 12th international conference on Knowledge-Based Intelligent Information and Engineering Systems, Part III*, KES '08, pages 67–75, Berlin, Heidelberg, 2008. Springer-Verlag.

[122] Thomas J. Schaefer. The complexity of satisfiability problems. In *Proceedings of the tenth annual ACM symposium on Theory of computing*, STOC '78, pages 216–226, New York, NY, USA, 1978. ACM.

[123] T. Schelling. Models of segregation. *American Economic Review*, 59(2):488–493, 1969.

[124] J. Scott and G. Marshall. *A Dictionary of Sociology.* Oxford Paperback Reference. Oxford University Press, 2009.

[125] Yaron Singer. How to win friends and influence people, truthfully: influence maximization mechanisms for social networks. In *Proceedings of the fifth ACM international conference on Web search and data mining*, WSDM '12, pages 733–742, New York, NY, USA, 2012. ACM.

[126] Jie Tang, Tiancheng Lou, and Jon M. Kleinberg. Inferring social ties across heterogenous networks. In *WSDM*, pages 743–752, 2012.

[127] Thomas W. Valente. *Network Models of the Diffusion of Innovations (Quantitative Methods in Communication Subseries).* Hampton Press (NJ), January 1995.

[128] Hal R Varian. A solution to the problem of externalities when agents are well-informed. *American Economic Review*, 84(5):1278–93, December 1994.

[129] Srinivasan Venkatramanan and Anurag Kumar. New insights from an analysis of social influence networks under the linear threshold model. *CoRR*, abs/1002.1335, 2010.

[130] Jan Vondrák, Chandra Chekuri, and Rico Zenklusen. Submodular function maximization via the multilinear relaxation and contention resolution schemes. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC '11, pages 783–792, New York, NY, USA, 2011. ACM.

[131] Wan-Shiou Yang, Jia-Ben Dia, Hung-Chi Cheng, and Hsing-Tzu Lin. Mining social networks for targeted advertising. In *System Sciences, 2006. HICSS '06. Proceedings of the 39th Annual Hawaii International Conference on*, volume 6, page 137a, jan. 2006.

[132] Pingzhong Tang, Tuomas Sandholm. Optimal iterative pricing over social networks. *Preprint*, AAAI-12, July, Toronto, Canada

[133] P. Maille, E. Markakis, M. Naldi, G. D. Stamoulis, B. Tuffin Sponsored Search Auctions : An Overview of Research with emphasis on Game Theoretic Aspects. *Working paper*

[134] Dimitris Fotakis, Paris Siminelakis. On the Efficiency of Influence-and-Exploit Strategies for Revenue Maximization under Positive Externalities. *Preprint*, arXiv:1110.1894v1 [cs.DS]

[135] David Kempe. Structure and Dynamics of Information in Networks. *Lecture notes.*

[136] Jennifer Wortman. Viral Marketing and the Diffusion of Trends on Social Networks. *Technical Report*, No. MS-CIS-08-19. University of Pennsylvania