UNIVERSITY OF TRENTO

MASTERS THESIS

CONSTRUCTIVE SOCIAL CHOICE WITH SETWISE MAX-MARGIN

Supervisor:

Prof. Andrea Passerini

Co-Supervisor:

Dr. Stefano Teso

Author:

Bereket Abera Yilma

A thesis submitted in fulfillment of the requirements for the degree of Master of Science in Computer Science

Department of Information Engineering and Computer Science

December, 2021

Declaration of Authorship

I **Bereket Abera Yilma**, hereby declare that this thesis titled ***CONSTRUCTIVE SOCIAL CHOICE WITH SETWISE MAX-MARGIN*** and the work presented in it are my own carried out under the supervision of ***Prof. Andrea Passerini and Dr. Stefano Teso***. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Sign:

Name & Reg. No.:

Date:

Abstract

In our day-to-day lives despite people possess conflicting preferences or attitudes they frequently arrive at a saddle point where they have to make joint (or collective) decisions. If not regularly often everyone participates in collective(group) decision making processes ranging from routine decisions or low-stakes such as at which restaurant to have lunch, which movies to watch,..etc to serious (high-stakes) decisions like with which equipments to assemble products of a company, which policy to approve, which chairman to elect,...etc. Coming up with a robust mechanism to make best possible group decisions has been researched earlier than the 18th century. A formal study providing a mathematical framework to study and develop mechanisms for group decision making was formalized as Social choice theory in the 18th century which is a sub-field of economics and political science. Recently joint approaches in Artificial Intelligence and computer science have been making efforts to addresses the problem of social decision making by thoroughly analyzing the process, various protocols and their efficiency for making the right decision.

However classical social choice focuses neither on learning nor on recommendation but only on analyzing voting protocols, and no existing "social preference learner" is constructive, as a result these approaches are incapable of solving the problem of social choice when set of options are combinatorial (i.e the item to be recommended is a combination of sub parts like PC configurations, sets of exam schedules, home furnitures...etc). This work addresses the problem of social choice from a constructive recommendation perspective (i.e recommending/choosing an outcome or a decision for a group by jointly eliciting personal preferences of individuals in a group generating entirely novel instances and recommending an instance that maximizes some notion of consensus among the group). We adopt a Constructive preference elicitation with Setwise Max-Margin as the back bone of our Constructive Social elicitation approach. The Setwise Max-Margin preference elicitation has been proved to produce a set of "diverse" items that can be used to ask informative queries to a group in order to get their feedback and use this feedback to iteratively tune the learning and come up with best recommendations that maximize consensus (Utility) among the group. Experimental results have shown that the constructive social choice algorithm is efficient in making the best constructive recommendation for a social choice task even in complex scenarios like PC recommendation task with reasonable computational cost hence achieving the goal of making a first step towards extending constructive preference elicitation to social choice domains.

Acknowledgements

First and foremost I want to thank God for giving me the strength and health to finish my studies.

- I wish to express my sincere thanks to Prof. Andrea Passerini and Dr. Stefano Teso for providing me an opportunity to do my thesis project under their supervision and for helping me gain further knowledge, for inspiring me to discover and learn in doing research work, and for their continuous support, patience, motivation and tireless guidance all the way. I was fortunate to benefit immensely from their guidance and intellectual support.
- I would like to express my gratitude towards my family for the encouragement and support they have given me, without whom none of these would have been possible.
- I am highly indebted to University of Trento and Opera Universitaria for giving me the opportunity and support to study in this university.
- Last but not least my thanks goes to all my friends who helped me over the course of this work and throughout my study.

Contents

List of Figures

The fear of the Lord is the beginning of wisdom . . .
Dedicated to the source of absolute wisdom.

Chapter 1

Introduction

1.1 Background and Motivation

Social choice has been the subject of intense investigation with the abundance of inexpensive preference data facilitated by recommender systems, search, online commerce and social networks. The online settings facilitating the elicitation, assessing and estimation of preference data from user populations stretches the boundaries of social choice within computer science, Artificial Intelligence, and operations research. Typically collective(group) decision making processes range from routine decisions domains or low-stakes such as (at which restaurant to have lunch, which movies to watch,..etc.) to serious (high-stakes) decision domains like (with which equipments to assemble products of a company, which policy to approve, which chairman to elect,...etc).

In all domains for assisting a collective or group wise decision the aggregation of individual preferences is required to form a single consensus recommendation, placing all these domains squarely in the realm of social choice Problem. For instance in determining potential products for a target market, ideally an online retailer would segment its audience into a group of S users in a way that a single product is desirable to be proposed for all members of the group. Another ideal domain is the first page of results returned by search engines like Google for a specific query. Ideally pure personalization would return results taking into account individual user's specific preferences , however this is generally not possible because of data scarcity. Hence the small amount of information known about one user's preferences is aggregated with the (equally scarce) preference information about users similar to that user in order to determine the best results. As a result a single set of results is constructed to be recommended for a collection of users despite the fact that they might possess conflicting preferences. Since this is a consensus decision making problem it can also be viewed as a social choice problem to be explored spanning through several areas in machine learning, with in the subfield of rank learning such as preference elicitation, online learning, structured output prediction and recommendation systems.

Various other factors make these domains and many other similar domains and their related problems both interesting and novel to be viewed and explored from a social choice perspective.

The social choice problem is essentially the problem of choosing or recommending an item that maximizes some notion of societal consensus or satisfaction. Coming up with a robust method for making societal decision requires the knowledge of individual preferences and the importance of each member in the group or society. A number of researches have been done in this direction over the past years since the 18th century. The prominent approaches to this problem include classical versions of Arrow's Theorem, the Gibbard-Sattherthwaite Theorem, the Median Voter Theorem and the maximal set.

Even though these approaches are the foundations of social choice theory, and most works done after these theories tried to analyze different scenarios and approach the problem of social choice from different dimensions, Classical social choice focuses neither on learning nor on recommendation but only on analyzing voting protocols, and no existing "social preference learner" is constructive. meaning the existing social choice approaches are incapable of solving the problem of social choice when the set of options are combinatorial, which references a research direction that has remained unexplored in the realm of social choice involves the recommendation of items that can be configured from sub pats on the basis of the user preferences, formally known as *Constructive recommendation* task. In constructive recommendation the recommended objects are "built" or "assembled" from scratch to maximize the user satisfaction. Constructive recommendations include recommendation of many types of objects that can be assembled from their components, for instance the item to be recommended is a combination of sub parts like PC configurations, sets of exam schedules, home furnitures, recipes, travel plans or shared schedules...etc).

Object synthesis for constructive recommendation usually involves solving a constrained optimization problem over a very large (or even infinite) domain of feasible configurations. For this particular reason reason, many of the existing social choice approaches are not suitable in this setting, since they are merely based on the assumption of a finite and relatively small set of available objects.

The large or sometimes even infinite feasible configuration spaces together with the availability of possibly controversial individual preferences and different importance of individuals in a group or a society makes the problem of social choice harder to solve.

This thesis focuses on making a first step to extend constructive preference elicitation approaches in to social choice domains following an incremental flow from constructive preference elicitation, to preference aggregation then social choice (group decision making). Our focus is on proposing an adaptive Societal preference elicitation framework a typical strategy as in[3][4] taking constructive view on social preference elicitation, enlarging the scope of preference elicitation from the selection of preference alternatives for a given society by applying some preference aggregation rule among a set of candidates alternatives to the synthesis of entirely novel instances.

In this chapter we first present social choice theory and existing preference elicitation and aggregation mechanisms, a brief overview of Constructive recommendation followed by problem description, Objective of the work and related literatures are visited subsequently.

1.2 Social choice theory

Social choice theory is a theoretical framework for a group decision making by analyzing and combining individual preferences, opinions, interests or welfares of members of a given society in order to derive a social preference that represents this society or community. The social preference is to express the general will, the common good or some notion of consensus among the society. Social choice theory is not a single theory, but a cluster of models and results concerning the aggregation of individual member's inputs into collective outputs.

1.2.1 History of Social Choice theory

Scientists from different fields have been contributing in different ways for the development of social choice theory since the 18th century. The prominent works include Jean-Charles de Borda and Nicolas de Condorcet from the 18th century and in the 19th century Charles Dodgson set a ground for the foundation of social choice theory. In the 20th century social choice theory started to raise with the works of Amartya Sen, Kenneth Arrow and Duncan Black. Ever-since the influence of social choice theory has been extending to different fields of research including Mathematics, political science, philosophy, economics and recently Biology and computer science.

Condorcet was also among the founding figures of social choice theory In addition to Arrow and Borda. A voting system known as the Borda count was defended by Borada which is often seen as a major alternative to majority voting. Some of the modern debates on how to respond to Arrow's theorem mostly arise from the debate between Condorcet and Borda. A comparison of simple majority, qualified majority, and unanimity rules and an analysis of the structure of preferences that payed crucial role as a precursor to later advances in social choice theory were proposed by the German statesman and scholar Samuel von Pufendorf (1632–1694). The theory of proportional representation in the 19 century following the independent rediscovery of many of Condorcet's and Borda's insights by the British mathematician and clergyman Charles Dodgson (1832–1898). was also among the advances of social choice theory. The Scottish economist Duncan Black (1908–1991) played a vital role in brining Condorcet's, Borda's, and Dodgson's social-choice-theoretic ideas to the attention of the modern research community. several discoveries related to majority voting are also attributed to Black.[2]

1.3 Preference Elicitation

Due to the explosion of on-line information making decision has become very challenging and as a result building effective decision support systems for on-line interfaces has been getting considerable research attention. In addition to the explosion of information what makes decsion making challenging task is the incomplete initial user preference and user's cognitive and emotional limitations of information processing capability. The how of eliciting the preference of users accurately is the main concern of current decision support systems.

In supporting users with decision making it is very crucial to accurately model user preferences. Since user preferences are incomplete initially(no preference informations are available at the start of interaction) in order for a decision support system to help users as they work towards their goal of decision making, preference elicitation techniques must attempt to collect as much information of users' preferences as possible. In addition to initial incompleteness of user preferences,and user's cognitive and emotional limitations of information processing, Preferences have a tendency of changing in different contexts. Thus preference elicitation techniques must be able to discover hidden preferences, avoid preference reversals and support users in making trade offs as they are confronting with competing objectives. Eliciting the preferences of users lies at the core of interactive decision support and recommendation systems. more formally it is the problem of developing a decision support system that are capable of generating high quality recommendations to a user hence supporting decision making.

Decision and utility theory can be viewed as the theoretical basis of user preference models. The Multi-attribute utility theory is more concerned on the evaluation of choices or outcomes for a problem of decision making.
The assignment of values to a set of attribute variables X = $\{X_1,\ X_2,...,\ X_n\ \}$ define Outcomes. Attribute variables could be either discrete or continuous. The space of all possible outcomes which is the Cartesian product of $\Gamma = \{\ X_1 \times X_2 \times ... \times X_n\ \}$ represents the outcome space. A decision making problem considers a set out comes O that are contained by Γ. It is common to have a very large Γ. For a decision maker to make a decision based on O a ranking of all outcomes determined by preferences is required and this is known as preference relation. Typically the preference relation under decision making problem of certainty is induced by a real-valued function v(o) : O $\rightarrow$ R .
The decision maker's preferences on a particular outcome is reflected by the value function. A more complex function "utility function" is usually needed to evaluate the "utility" of a decision under cases of uncertain decision scenarios , where the outcomes are characterized by probabilities. The user's attitudes towards risk and the value of outcomes are what a utility function represents, which also induces a preference ordering

on the probability distributions over the outcome space. For a utility function u in order to correctly preserve the user's preference relation for actions it must consider the user's attitudes toward risk and the uncertainty of attaining outcome o, while assigning values for o. Various assumptions have been made concerning the structure of preferences to overcome the time-consuming and tedious nature of value (utility) function elicitation over large amount of outcomes. Additive independence is the most commonly applied assumption where utility(value) of an outcome is broken down to the sum of individual attributes. The reduction of a number of outcomes to be considered and the construction of more manageable and less complicated value functions are allowed by the assumption of independence. But since attributes are preferentially dependent in many cases the assumptions of decomposability are incorrect.

The elicitation of preferences of a new user using the structures of the closest existing preference as potential default in order to save user's effort as much as possible and elicit utility (full value) function has been proposed by some research works. These works assume that there exists a complete or incomplete structures of preference that are elicited from a population of users and no restrictive assumptions are made on the form of the underlying value functions.

1.4 Constructive recommendation

Constructive recommendation is the task of recommending structured objects, i.e. configurations of several components, assembled on the basis of the user preferences (Teso, Passerini, and Viappiani 2016; Dragone et al. 2016). The recommended objects are "built" or "assembled" from scratch out of exponentially growing space of feasible configuration spaces to maximize the user satisfaction. Recommendations include many types of objects that can be assembled from their components, for instance personal computers or mobile phone configurations, sets of exam schedules, home furnitures, recipes, and complex preference-based decision problems, such as customized travel planning or shared personalized activity scheduling.[16]

Preferences play vital role role not only in recommendation systems but also in a variety of artificial intelligence applications and the task of eliciting or learning preferences is a crucial strategy on the way to making recommendation ; typically only limited information about the user's preferences will be available and the cost (cognitive or computational) of obtaining additional preference information will be high. Tackling this by solving a constrained optimization problem is the main task of constructive recommendation. Constructive recommendation takes a *constructive* view on preference elicitation, enlarging its scope from the selection of items among a set of candidates to the synthesis of entirely novel Configurations. *Configurations* are solutions to a given optimization problem; they are represented as combinations of basic building blocks or sub parts (e.g. the components of a computer) subject to a set of constraints (e.g. the laptop model determines the set of available CPUs).

Each configuration or object is represented as a set of attributes, e.g. brand, shape, size, color,...etc. i.e. Boolean or numeric variables that describe the physical properties of the objects. The space of possible object configurations in the number of attributes is exponential or even infinite if some attributes are continuous variables. The possible configurations can also be constrained by arbitrary boolean formulas to limit the feasible space. A utility function is learned over the feature representation of an instance, as customary in many preference elicitation approaches. The recommendation is then made by solving a constrained optimization problem in the space of feasible instances, guided by the learned utility.[3]. (i.e Constructive recommendation consider the problem of finding the best object for the user in this large combinatorial space of candidates.)

1.5 Problem description

In our day-to-day lives despite the fact that people possessing conflicting preferences or attitudes they frequently arrive at a saddle point where they have to make joint (or collective) decisions. If not regularly often everyone participates in collective(group) decision making processes ranging from routine decisions or low-stakes such as (at which restaurant to have lunch, which movies to watch,..etc.) to serious (high-stakes) decisions like (with which equipments to assemble products of a company, which policy to approve, which chairman to elect,...etc.). coming up with a robust mechanism to make best possible group decisions has been researched earlier than the 18th century. Most works in this domain represent people's preferences as utilities or cardinal preferences in which a user's preference is specified by some utility function U.

Utilities are numerical values that are associated with the available alternatives. For any preference ordering it is common to find a set of consistent utilities not violating the underlying ordering. preferences either ordinal or Cardinal each offer certain advantages over the other. A utility function encodes more information about individuals' preferences (i.e., it is more expressive), including the intensity of preferences and representing preferences in the form of a probability distribution nonetheless, in the absence of common numeric value such as money, interpersonal comparison between utilities is quite problematic and controversial. On the other hand arriving at the best consensus decision requires efficient way of eliciting individual preferences.

The problem of social choice has not yet been addressed from a constructive recommendation perspective where the items recommended are generated from scratch instead of being chosen from among a set of available options. In a social choice problem when the set of options are combinatorial (i.e the item to be recommended is a combination of sub parts like PC configurations, sets of exam schedules, home furnitures...etc) robust constructive recommendation approach is needed in order to make the best decision for the group or society.

These together with the fact that much of the social choice literature deals with neither learning(utilities) nor on recommendation but merely focus on rankings and most works assume that items or decisions are available in a (possible large) dataset—motivates us to focus on Constructive social choice following an adaptive preference elicitation framework that takes a constructive view on preference elicitation in this thesis for solving the social choice problem in combinatorial alternative domains.

1.6 Related work

Social choice addresses the problem of choosing an outcome or a decision, for a society of individuals (or agents) who have their own personal preferences over the set of alternatives. Different efforts tried to address the social choice problem in the past including Probabilistic and Utility-theoretic Models in Social Choice. In most social choice problems it is unlikely to find complete preference expression because users simply do not tolerate much in the way of elicitation. Therefore preferences are usually estimated from choosing behavior, partial rating data,..etc. The learning of quite compelling probabilistic models of user preferences could be made feasible by having large amount of such data. Probabilistic model, refers to some distribution P over the set of preferences or rankings V . For several reasons approximation is an absolute necessity. For instance computational approximation is demanded for "nearly instantaneous" recommendations, the design of approximation methods greatly facilitate to economic tradeoffs (informational approximation fills the gap in the incompleteness of preference data).

As customary in many preference elicitation approaches; in this social elicitation approach an aggregate utility function is learned over the feature representation of instances. The recommendation for a group or a society is then made by solving a constrained optimization problem in the space of feasible instances, guided by the learned aggregate utility. Different elicitation methods have been proposed previously to tackle Preference elicitation in configuration problems. One approach was regret-based elicitation which uses minmax regret as a technique to drive elicitation and also as a robust recommendation criterion[Boutilier et al., 2006; Braziunas and Boutilier, 2007]. This was limited by the lack of tolerance with respect to user inconsistency. Since dealing with uncertain and possibly inconsistent user feedback in learning a user utility function is a key requirement, the lack of tolerance with respect to user inconsistency is considered the main limitation of regret-based elicitation. Experimental results from [4][5] proved that Bayesian preference elicitation approaches to this problem are computationally expensive and can not scale to fully constructive scenarios.[Chajewska et al., 2000; Guo and Sanner, 2010; Viappiani and Boutilier, 2010].These approaches tried to solve the problem by an interactive way asking queries maximizing informativeness measures like expected value of information (EVOI) and building a probability distribution on candidate functions (endowed with a response or error model to be used for inference).

Extensive social choice studies have been done with issues of computational approximation and informational approximation that deal with incomplete preference information mostly resorting probabilistic models.When tackling various problems in social choice primarily realistic models that can be effectively learned from readily available data and support tractable inference are needed. In order to model population preferences, models that explicitly try to reflect the processes of comparison judgments by human have been developed in statistics, econometrics and psychometrics.[5]. Under the guise of "learning to rank" (LeToR) the machine learning community has appropriated several of these models especially the Mallows and Plackett-Luce models. As a result precipitating the development of several interesting approaches for probabilistic inference and tractable learning with such models.

The development of new algorithms and models for probabilistic inference. and learning is being influenced by the problems that arise in social choice. For instance observed rankings are assumed to be noisy estimates of some underlying objective ranking (rather than representing genuinely distinct preferences)by most works in in LeToR.[6] Several important problems have been unaddressable due to the types of data sets considered. Learning Mallows models can be seen as one example which was widely considered intractable with choice data consisting of pairwise comparisons of the form $x \succ x'$ meaning x is preferred tox', conspicuously an important form of evidence in any social choice problem. The generalized repeated insertion model (GRIM) is one of the models developed allowing Mallows models (and mixtures thereof) to be effectively learned from such data. GRIM allows approximate sampling of rankings conditioned on pairwise evidence and interesting population models have been learned with various real world data sets employing this technique[7]. Models that are general enough to address tractability and support effective inference are needed. Especially consensus decision making domains expect to find models that are tuned to the types of preference distributions within specific domains of interest. Seemingly unaddressed; tractable and realistic models for distributions over single-peaked preferences could be mentioned as an example. ,

In the development of social choice methods another key issue is centered on the notion of utility theoretic approximation of recommendations or "winners", especially while having partial preference information about users. Incomplete preferences have been studied in a variety of cases but unfortunately the question of how to select a winner in such a situation was given little attention. The notion of minimax regret (MMR) has been proposed to tackle this issue[8]. As customary in most works on this domain MMR employs voting rules defined using some natural *scoring function* $f(x;\mathrm{H})$ which measures the utility or quality of preference alternative x given profile H. (i.e γ (H) $\in \mathrm{argmax}_{x\in X}$) $f(x;\mathrm{H})$. Having access only to partial votes of some of the voters; (i.e., replacing each vote h with a (possibly empty) partial order p, or a collection of pairwise comparisons and P denoting this *partial profile*; the MMR approach tries addresses the problem of

selecting a winner intuitively. measuring the quality of x given p by considering how far from optimal x could be in the worst case (i.e., given any *Completion* or extension H $\in$ C(P) of p). The minimax optimal solution is any preference alternative that is nearest to optimal in the worst case.
MMR can be formally presented as:

$$Regret(x, \mathrm{H}) = max_{x' \in X}\ f(x'; \mathrm{H}) - f(x; \mathrm{H})$$
$$MR(x, \mathrm{P}) = max_{H \in C(P)}\ Regret(x, \mathrm{H})$$
$$MMR(P) = min_{x \in X}\ MR(x, \mathrm{P})$$
$$\hat{x_p*} \in argmin_{x \in X}\ MR(x, \mathrm{P})$$

The tightest possible bound on loss of societal utility is provided by the minimax winner $\hat{x_p*}$, and for a variety of voting rules MMR can be computed in Polynomial time and provides distinct recommendations analogized with selecting among possible winners. However MMR fails to exploit distributional information P about voter preferences. With such a probabilistic model it is possible to select a winner by maximizing the value of expected Utility(**MEU**): $\hat{x_p*} = argmax \sum_H P(\mathrm{H}|\mathrm{P})\ f(x; \mathrm{H})$.

For various combinations of voting rules and preference distributions this has been computationally challenging problem to solve. Despite ensuring (Bayesian) optimality in the presence of a partial profile, MEU provides no guidance w.r.t. potential loss relative to choosing a winner with a complete voting profile H. In contrast MMR tells the potential value of adding new evidence to complete a voting profile through *expected regret(ER)* regarded as the most natural way of seeing loss with respect to a proposed alternative x.
$ER(x, \mathrm{P}) = \sum_H P(\mathrm{H}|\mathrm{P})\ Regret(x, \mathrm{H})$
Another critical process in social choice is elicitation of preferences/votes. The same alternative $\hat{x_p*}$ maximizes expected utility and minimizes expected regret, but ER is regarded as more useful and informative for elicitation purposes. MMR could be most natural criterion for robust selection of alternatives in the absence of a probabilistic model P, however if MMR is too big the potential error associated with any winner will be unacceptable. In order to reduce MMR an interactive approach to elicitation was suggested (i.e, by asking voters some queries about their preferences[8]. In general Bayesian approaches to preference elicitation build a probability distribution on candidate functions bestowed with error or a response model to be used as an inference and ask queries in order to maximize informativeness measures such as MEU, ER and expected value of information (EVOI). However these approaches have been proven to

be computationally expensive and can not scale to fully *constructive scenarios* (Teso, Passerini, and Viappiani 2016) where the task is synthesizing entirely novel instances to be chosen that satisfy the some notion of consensus within a society, like the furniture arrangement of an apartment or a PC configuration to be bought by one family..etc. Figure 1.1 shows the comparison between these approaches.
The comparison demonstrates the recent approach SETMARGIN can easily scale to much larger datasets and constructive scenarios compared with three state of the art Bayesian approaches and reports solution quality and timing values for increasing number of collected user responses.

1. the Bayesian approach from [Guo and Sanner, 2010], selecting queries according to restricted informed VOI (RIVOI)

2. the Bayesian framework of [Viappiani and Boutilier, 2010] that uses a greedy optimization of Expected Utility of a choice (a tight approximation of EVOI, hereafter just called EUS).

3. Query Iteration (referred as QI below), also from [Viappiani and Boutilier, 2010]

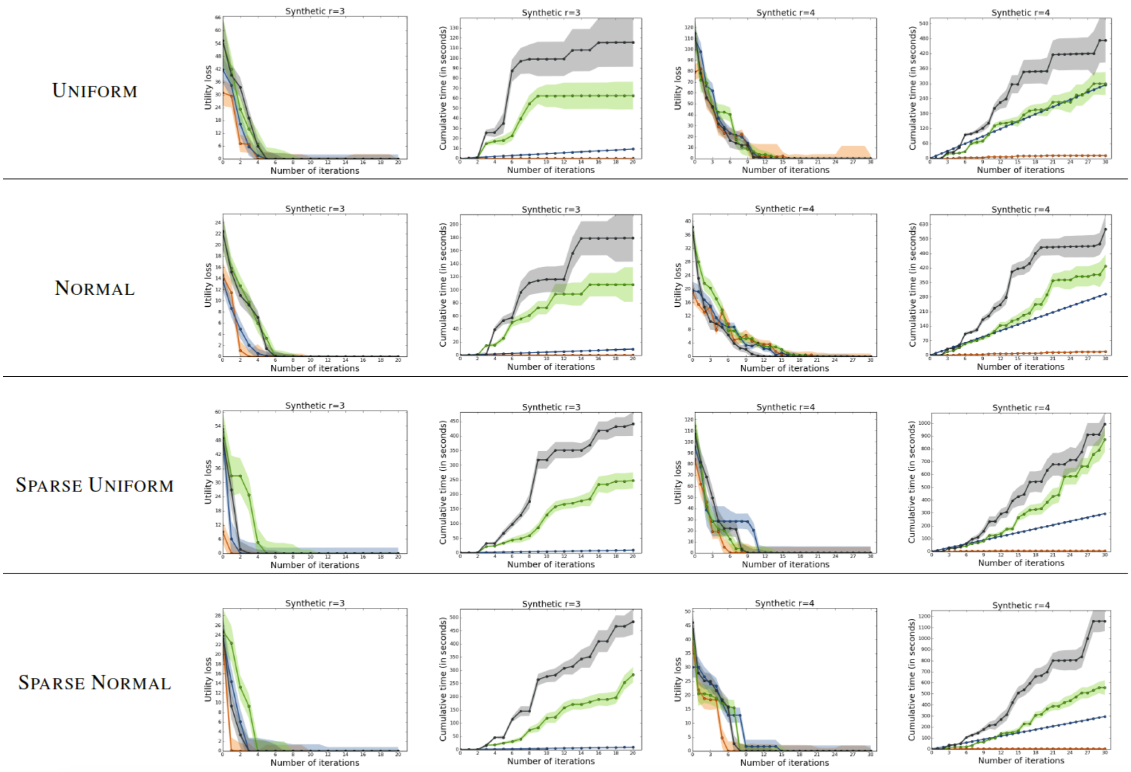


Figure 1.1: Comparison between SETMARGIN (orange), RIVOI (blue), QI (green) and EUS (gray) on the (attributes)r = 3 (left) and r = 4 (right) datasets.

Each row represents a different sampling distribution for user utility. The number of iterations is plotted against the utility loss (first and third columns) and the cumulative time (second and fourth columns). Thick lines indicate median values over users, while standard deviations are shown as shaded areas. (Source: https://arxiv.org/pdf/1604.06020.pdf)

1.7 Objective

At the heart of social choice problem or group decision making lies preference elicitation. As proposed by a number of researchers a most natural approach is asking informative questions which often makes it possible to make near-optimal decisions with only partial preference information. However most of these works assume decisions or items are available in a possibly larger datasets and they build a probability distribution on candidate functions bestowed with error or a response model to be used as an inference and ask queries in order to maximize informativeness measures such as MEU, ER and expected value of information (EVOI) which are proven to be computationally expensive and can not scale to fully constructive scenarios.

Classical social choice focuses neither on learning nor on recommendation but only on analyzing voting protocols, and no existing "social preference learner" is constructive. meaning these approaches are incapable of solving the problem of social choice when the set of options are combinatorial(i.e the item to be recommended is a combination of sub parts like PC configurations, sets of exam schedules, home furnitures...etc).

Since constructive learning in social choice is a natural extension and many problems can be measured in it; in this thesis we set the goal of making a first step towards extending constructive preference elicitation to social choice domains where a single best item is recommended to a group of people or a society for instance establishing a shared exam schedule among professors at some university, arranging furnitures of a family home or buying a PC for a whole family. In analyzing social choice problems the tricky part is that users or individual members in a society might have different preferences (utilities) and another important aspect not to forget is that they also posses different importance or saying for instance we can think of the head of the department in the scenario of deciding a shared exam schedule, or the parents in a family who actually pay for the PC and get the last say on it.

In this thesis we propose an adaptive elicitation frame work employing SETWISE MAX-MARGIN elicitation techniques which takes a constructive view on preference elicitation, enlarging its scope from the selection of preference alternatives for a given society by applying some preference aggregation rule among a set of candidates alternatives to the synthesis of entirely novel *Instances*. By *Instances* we refer to decisions or items to be chosen/recommended (i.e solutions that are diverse and can efficiently maximize the societal utility represented as combination of basic elements that are subject to a set of constraints).

Chapter 2

Methodology

2.1 Constructive Social Choice

The problem of social choice has not yet been addressed from a constructive recommendation perspective where the items recommended are generated from scratch instead of being chosen from among a set of available options. In a social choice problem when the set of options are combinatorial (i.e the item to be recommended is a combination of sub parts like PC configurations, sets of exam schedules, home furnitures...etc) robust constructive recommendation approach is needed in order to make the best decision for the group or society. While most works of Social choice and group decision making assume that items or decisions are available in a (possible large) dataset. In this work we propose an adaptive elicitation frame work that takes a *constructive* view on social preference elicitation, enlarging its scope from the selection of preference alternatives for a given society by applying some preference aggregation rule among a set of candidates alternatives to the synthesis of entirely novel instances. By *Instances* we refer to solutions to a given optimization problem that can be represented as combinations of basic elements that are subject to a set of constraints. for instance considering the components of a laptop, the laptop model determines the set of available CPUs. In social choice or group decision making weather the set of available options are combinatorial or not having a prior knowledge about the individual user's preferences and their corresponding importance is crucial and therefore making *preference elicitation* a first step in social choice.

This work tackles Social elicitation in configuration problems inspired by previous works of preference elicitation with setwise Max-Margin learning method [4] [5]. We take a space decomposition perspective and jointly learn a set of parameters $\omega \in R^M$ each of which representing a candidate aggregate utility function for a group of M users or a society maximizing the diversity between the instances and the consistency of the available feedback. This approach to social elicitation works by combining instance generation with weight vector learning, meaning each iteration of the algorithm yields two outcomes, a set of weight vectors representing individual users contribution in a group and a set of instances(configurations) each maximizing the group-wise utility according to individual user's contribution. The Setwise Max-Margin method used for social elicitation will be discussed in the following section

2.2 Setwise Max-Margin

As a customary to Social choice approaches elicitation task unfolds the aggregation of individual user preferences. A bottom up discussion is presented in subsequent sections from individual preference elicitation to preference aggregation and social choice.

Setwise max-margin learning extends the goal of max-margin which is establishing a singleton solution that is as robust as possible to producing a set of solutions instead. In several machine learning algorithms the notion of margin is very crucial. A learner ensures robustness in it's solution by maximizing the margin. Results from the extension of max margin learning are used to generate a set of recommendations and diverse set of feasible utility functions that are useful for further elicitation in asking users a choice query based on this set ("Among these list of items, which one do you prefer ?").

Notations to be used in subsequent sections

- Boldface letters $\boldsymbol{x}$ to indicate vectors,
- Uppercase letters X represent matrices,
- Uppercase letters *X'* represent matrix transposition.
- Calligraphic capital letters $\mathcal{X}$ for sets.
- $||X||_1 := \sum_z |X_z|$ indicate ℓ_1 vector norm.
- $\langle . \ , . \rangle$represent the usual dot product
- x= $(x_1, \ . \ . \ . \ ,x_N)$ represents configurations over N feature space (N attributes).
- $X_z \in \{0{,}1\}$ for all $z\in$ [n] assuming a one-hot encoding of categorical features.
- $w \in R^N$weight vectors representing user's preferences
- $\langle x,w \rangle = \sum_{z=1}^{n} w_z x_z$ for utility of a configuration x

- All weights are required to be non-negative and bounded $[w_z^{\perp},w_z^{\top}]$ with $w_z^{\perp} \geq 0$.
- A set of users $u \in \{1,....M\}$
- $w_u^{*}\in R_{+}^{N}$ representing true preferences of user $u \in [M]$
- $x_u^{*} \in \mathcal{X}$ representing one of the configurations most preferred by user u.
- $K \in N$ represents Cardinality of the query sets
- $w_1^u, \ . \ . \ . \ ,w_K^u \in R_{+}^{N}$ (learned)estimated preferences of user u.
- $x_1^u, \ . \ . \ . \ ,x_K^u \in \mathcal{X}$ Query set made to user u.
- $x^u\in \mathcal{X}$ recommendation made to user u.
- $v(u) \geq 0$ variability within $\{w_i^u\}$.
- $k(u,y) \geq 0$ Similarity between $\{w_i^u\}$ and $\{w_i^y\}$.
- $\delta := (\alpha,\beta,\gamma) \in R_{+}^{3}$ Hyper parameters of the mu-swmm algorithm.
- $W \in R_{+}^{N\times M}$ is the matrix obtained by stacking the learned preference vectors of all users side-by-side(aggregation).
- $W_{*} \in R_{+}^{N\times M}$ is the matrix obtained by stacking the true preference vectors of all users side-by-side(aggregation).
- $\omega\in R^{M}$ the contribution of different users to the overall utility of configurations.
- $\mathcal{D}$ represents the set of preferences (answers to comparison queries) elicited from users.

The non- negativity and and boundedness of the weight vectors enables the translation of the core optimization problem into a mixed-integer linear problem(MILP).

The setwise max-margin methodology: the actual weight vector w is unknown to the learning system while learning, and therefore must be estimated by interacting with the user. The interaction focuses on asking pairwise comparison queries, which are assumed to be the simplest of the comparative queries and this can be extended to choice sets of more than two alternative options. For a pairwise comparison between two configurations x and x' either of the following conditions are true. x is preferred to$x$' (written as $x \succ x$'), x' is preferred to x ($x' \succ x$) or there is no clear preference between the two items ($x \approx x$') Setwise Max-Margin approach has two fold objectives. The first learning objective is finding a set of k weight vectors $w^1, \ . \ . \ . \ , w^K$ that are chosen so that all the preferences provided by a user are satisfied by largest possible separation margin and will be maximally diverse. The second objective is constructing a set of k configurations $x^1, \ldots, x^K$ that are maximally diverse among each other and each of the generated configurations x^i are the best possible options when evaluated according to their corresponding w^i

The first objective is achieved by getting ranking constraints translated from all pairwise preferences ($\mathcal{D}$). Pairwise preferences of the form $y_+^h \succ y_-^h$ become linear inequalities $\langle w^i, y_+^h - y_-^h \rangle \geq \mu$. where μ represents the *margin variable* to be maximized and h ranges over the responses. Occasional inconsistencies of user feedback in practice usually resulting non separable dataset is handled by introducing slack variables (whose sum is aimed to be minimized).

The above translated inequality augmented with slacks takes the following form:
$\langle w^i, y_+^h - y_-^h \rangle \geq \mu - \xi_h^i$. where ξ_h^i stands for the penalty incurred by w^i for violating the margin separation of pair h.
The second objective is achieved by maximizing the sum of utilities $\sum_{i=1}^{k} \langle w_i, x_i \rangle$ and adding ranking constraints of the form $\langle w^i, y_+^h - y_-^h \rangle \geq \mu$, $\forall\ i,j \in [\mathcal{K}]$, $i \neq j$.
Setwise Max-margin choses the configurations $\{x^i\}$ and the weight vectors $\{w^i\}$ simultaneously meaning that the algorithm looks for w^i so that the utility loss of choosing x^j instead of x^i, $j \neq i$ is large at least μ (the margin).
Figure 2.1 is a demonstration of this formulation choosing a pair $k = 2$.
The black lines represent preference constraints $\mathcal{D}$, the red points represent the utility

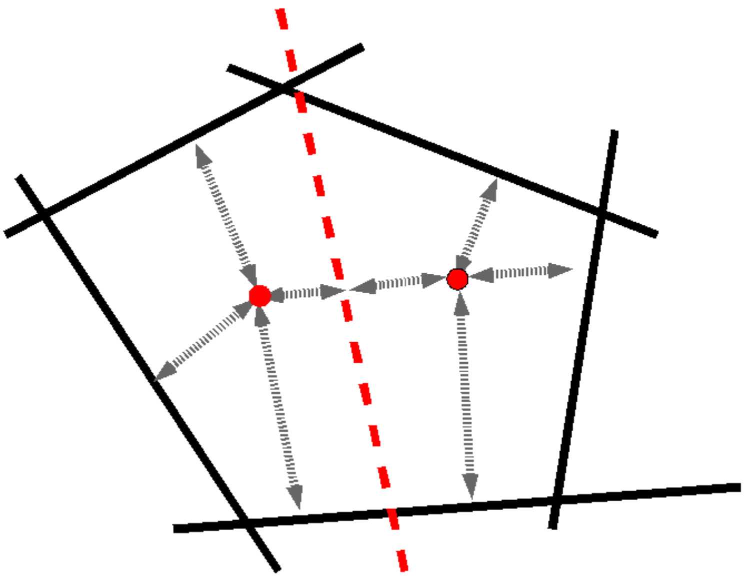


FIGURE 2.1: Optimization of setwise max-margin;

vectors w^1 and w^2, and the red line represents the hyperplane $\langle w, x^1 - x^2\rangle = 0$ partitioning the space of feasible utility weights in two parts (in general, there will be k subregions).

2.2.1 Multi-user Setwise max-margin

To make a group decision or a general recommendation to a society containing multiple users as one, preference aggregation requires to know or elicit individual user's preferences in advance. In this work we adopt constructive preference elicitation strategy for multiple users with Setwise Max-margin to learn the individual user preferences simultaneously by an interactive way as in [4]. This is also a realistic scenario of real systems such as electronic commerce websites that do not usually interact with single user in isolation, but are accessed by several users at the same time hence simultaneously eliciting the preferences of multiple users. we focuses on constructive recommendation tasks. Here setwise max-margin optimization method, is used which is viewed as a generalization of max-margin learning to sets, that supports the identification of informative questions and encouraging sparsity in the parameter space.

In this method the instance to be recommended is synthesized by searching in a constrained configuration space rather than choosing among a set of pre-determined preference alternatives. In every stage of the interaction process individual users are associated with parameter weights (that are viewed as alternative options for the unknown user utility) and this is used to identify users that are similar and to propagate preference information between them. The multi-user setwise Max-Margin(MUSM) focuses on constructive recommendation problems in a case where a number of users are simultaneously present. And the task is to generate novel configurations that are subject to user preferences and feasibility constraints instead of selecting an item among a set of candidate alternatives. This rues out a standard method where recommendations propagate between users based on shared ratings over similar items by collaborative filtering. MUSM rather rely on a notion of similarity in model space, i.e. similar users have similar utility functions, hence propagate preference information among similar users while simultaneously learning user utilities.
The MUSM elicitation is achieved by

- Defining a user similarity as utility models;
- Measuring the reliability of the learned model for each user;

- Defining an aggregate utility function for each user combining his/her utility model with those of the other users, weighted by their respective reliability and similarity to the user being recommended.

MUSM incorporates these aspects in the set-wise max-margin optimization problem, and retains the formulation as mixed integer-linear problem (MILP) which allows for efficient computation. Simultaneously eliciting preferences of multiple users, MUSM distributes the cognitive cost of query answering over multiple users and also exploits the similarity between users in order to transfer the gained knowledge among them. Experimental results has proved that MUSM provides high utility recommendations to individual users by simultaneously eliciting the preferences of different users considering all available preference information. Here we focus on extending constructive preference elicitation to social choice domains employing MUSM as individual user's preference elicitation tool.

2.3 Social Elicitation

Having the preference information of individual users in a society or a group, the analysis of constructive social choice recommendation task follows preference aggregation which is at the heart of social choice theory commonly understood as aggregation of several individuals' preference rankings of two or more social alternatives into a single, collective choice ranking (or preference). The Constructive social choice algorithm works by assuming users in a society or a group have leaned preference information (eg, using multi-user setmargin $w_u \in R^{*}_{+}$) or assuming true preference of users ($w^{*}_{u} \in R^{*}_{+}$) are available in advance. The aggregate utility matrix W obtained by stacking the learned individual preference vectors side by side takes the form $W \in R^{N\times M}_{+}$. And $W_{*} \in R^{N\times M}_{+}$ is the analogue for the true preferences.

The aggregate Preference of the users as a group are defined as a function f of W. Since linear combinations can capture the contribution of different users to the overall utility of configurations. we will focus on linear combinations:

$$f(x;W) := \sum_{u=1}^{M} \omega_u \langle w_u, x \rangle$$

Unlike MUSM, in the Social elicitation scenario the interaction is with a group collectively not to individual users. Here the goal is two fold similar to that of the Setwise max-margin preference elicitation strategy. The first objective is to learn the parameters $\omega \in R^M$ representing the contribution of different users to the overall utility of configurations from collective, group-wise feedback by an interactive way as in MUSM preference

elicitation methods. The second objective is Selecting a set of S best informative configurations $x^1, \ldots, x^S$ with maximal aggregate utility and diversity. The interaction with the group focuses on asking pairwise comparison queries and this query set S is used to make questions to the group of users as a whole and collect feed back of the form $x^1 \succ x^2$. This requires solving an optimization problem.
In a similar manner as set margin preference elicitation strategy for a pairwise comparison between two configurations x and x' either of the following conditions are true. x is preferred to x' (written as $x \succ x'$), x' is preferred to x ($x' \succ x$) or there is no clear preference between the two items ($x \approx x'$).
The first learning objective is achieved by getting ranking constraints translated from all pairwise preferences $\mathcal{D}$ (collective group-wise feedback).
When queried with a set of S configurations $x^1, \ldots, x^S$, the group indicates the overall most preferred configuration x^i. This group wise feedback is an aggregate of individual users contribution.
The learning is done by iteratively updating the user contributions ω to f using the group wise feedback. This means users ranking x^i (a configuration in the query set) high are most important and users ranking x^i low are considered less important in learning the optimal ω for the society or the group. This means finding an ω which minimizes the ranking loss over the dataset $\mathcal{D}$(ranking constraints that it satisfies).

- Feedback: choosing x^i out of the S query configurations amounts to a set of S - 1 ranking constraints of the form:

$$\forall\, j \in [S], j \neq i;\, f(x^i) \succ f(x^j)$$

$$\Leftrightarrow \langle \textstyle\sum_{u=1}^{M} \omega_u w_u, x^i \rangle \geq \langle \textstyle\sum_{u=1}^{M} \omega_u w_u, x^j \rangle$$

$$\Leftrightarrow \textstyle\sum_{u=1}^{M} \omega_u \langle w^u, x^i \rangle \geq \textstyle\sum_{u=1}^{M} \omega_u \langle w^u, x^j \rangle$$

$$\Leftrightarrow \textstyle\sum_{u=1}^{M} \omega_u \langle w^u, x^i - x^j \rangle \geq 0$$

The gradient of this expression w.r.t. ω^u for a single pair i,j is:

$$\frac{\partial}{\partial \omega_u} \textstyle\sum_u \omega_u \langle w^u, x^i - x^j \rangle = \langle w^u, x^i - x^j \rangle$$

This leads to a Perceptron update of the form:

$$\omega_u{}^{t+1} \leftarrow \omega_u{}^{t} \pm \eta \textstyle\sum_{j \neq i} \langle w^u, x^i - x^j \rangle$$

More concisely:

$$\omega_u{}^{t+1} \leftarrow \omega_u{}^{t} \pm \eta W^{\top} \delta$$

Where η represents the step size (learning rate $\frac{2}{S(S-1)}$) and $\delta_{j \neq i} = (x^i - x^j)$

Selecting a query : The objective of selecting best informative queries to present to the group in order to get their feedback is achieved by essentially solving an optimization problem which requires maximizing the sum of aggregate utility(sum over the utilities of all objects in the query set) $\sum_{u=1}^{M} \omega_u \langle w_u, x \rangle$ and the diversity of the generated configurations, This can be formalized as the following:
Choosing a pair S = 2 for simplicity.

$$\max_{x^1,..,x^2} \lambda \sum_{i=1}^{S} f(x^i) + (1 - \lambda) \sum_{i<j} ||x^i - x^j||_1$$

$$\text{s.t } f(x^i) \neq f(x^j) \text{ , } \forall\ i,j \in [S], \text{ i<j}$$

The first term maximizes the sum of aggregate utilities, sum over the utilities of all objects in the query set (i.e the query set to be high quality) and the second term maximizes the diversity of the selected configurations. In order to combine these two parts we used a simple convex combination λ for the first part and (1 - λ) for the second part. where $\lambda \in$ [0,1] is a user-provided hyper-parameter. and $||.||_1$ is the ℓ_1 norm,

$$\text{i.e } ||x||_1 := \sum_{z=1}^{N} |x_z|$$

Maximizing the ℓ_1 norm over a categorical representation is equivalent to requiring that the configurations $x^1, \ . \ . \ . \ , x^S$ have a maximal number of different attributes; (i.e if x^1 and x^2 differ in all attributes,regardless of the actual values, they achieve maximal ℓ_1 distance.)

A straightforward encoding of the initial formulation is problematic to encode as (mixed integer linear program)MILP. The derivation for suitable encoding as MILP could be done as follows:
The case for $S = 2$ for simplicity.

$$\max_{x^1,x^2} \lambda(f(x^1) + f(x^2)) + (1 - \lambda)\ ||x^i - x^j||_1$$
$$\text{s.t } f(x^1) \neq f(x^2)$$

$$= \max_{x^1,x^2} \lambda(f(x^1) + f(x^2)) + (1 - \lambda) \sum_{z=1}^{N} |{x_z}^1 - {x_z}^2|$$
$$\text{s.t } f(x^1) \neq f(x^2)$$

$$= \max_{x^1,x^2,\epsilon} \lambda(f(x^1) + f(x^2)) + (1 - \lambda)\sum_{z=1}^{N} \epsilon_z$$
$$\text{s.t } \epsilon_z \leq |{x_z}^1 - {x_z}^2|, \forall z = 1,.......N$$
$$f(x^1) \neq f(x^2)$$

$$= \max_{x^1,x^2,\epsilon} \lambda(f(x^1) + f(x^2)) + (1 - \lambda)\sum_{z=1}^{N} \epsilon_z$$
$$\text{s.t } \epsilon_z \leq \max \left\{{x_z}^1 - {x_z}^2, {x_z}^2 - {x_z}^1\right\}, \forall z = 1,.......N$$
$$f(x^1) \neq f(x^2)$$

$$= \max_{x^1,x^2,b^1,b^2,\epsilon} \lambda(f(x^1) + f(x^2)) + (1 - \lambda)\sum_{z=1}^{N} \epsilon_z$$
$$\text{s.t } \epsilon_z \leq ({x_z}^1 - {x_z}^2) - {b_z}^1.M \ , \forall z = 1,.......N$$
$$\epsilon_z \leq ({x_z}^1 - {x_z}^2) - {b_z}^2.M, \forall z = 1,.......N$$
$${b_z}^1 + {b_z}^2, = 1, \forall z = 1,.......N$$
$$f(x^1) \neq f(x^2)$$

The *big-M* trick:
Here ϵ and b are variables introduced in order to linearize the problem.
M is a large positive constant,(e.g, $M = 10^6$).

Individual preference information elicitation is not restricted to multi-user set margin. It could also be collected using other mechanisms as well but from the perspective of the Social elicitation optimization problem nothing changes.

Algorithm 1 The CONSTRUCTIVE-SOCIALCHOICE algorithm; it uses the MUSETMARGIN algorithm as first step to receive the individual user information. Here M is the number of users, S is the set size, and T is the number of iterations. The hyperparameter λ is left implicit.

1: **procedure** CONSTRUCTIVE-SOCIALCHOICE(m, w^u, S, T)
2: ; Initialization
3: **for** $u = 1, \ldots, m$ **do**
4: $MUSM(MultiUserSetMargin)$
5: ; Main loop
6: $\mathcal{D}_u \leftarrow \varnothing$
7: **for** $t = 1, \ldots, T$ **do**
8: $\{\omega^i, \boldsymbol{x}^i\}_{i=1}^S \leftarrow$ SOLVE$(\mathcal{D}, S, \lambda)$
9: **for** $\boldsymbol{x}^i, \boldsymbol{x}^j \in \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^m\}$ **s.t.** $i < j$ **do**
10: THE GROUP SELECTS$(\boldsymbol{x}^+ \text{from} \{\boldsymbol{x}^i\})$
11: $\mathcal{D} \leftarrow \mathcal{D} \cup$ QUERYGROUP$(\boldsymbol{x}^i, \boldsymbol{x}^j)$
12: $\omega^*, \boldsymbol{x}^* \leftarrow$ SOLVE$(\mathcal{D}, 1, \lambda)$
13: **return** $\omega^*, \boldsymbol{x}^*$

The Constructive social choice procedure is sketched in Algorithm 1. At every iteration, the algorithm selects S query configurations $x^1, \ . \ . \ . \ , x^S$ based on the previously collected group wise feedback $\mathcal{D}$ (line 8). The query set $\{x^i\}$ is presented to the group, invited to select a most preferred configuration x$^+$ from the S alternatives (line 10). The group choice is interpreted as a set of pairwise ranking constraints $\{x^+ \succ x^- : x^-$ was not selected $\}$ and added to $\mathcal{D}$ (line 11). At this point, a recommendation x is computed by leveraging all group-wise feedback and presented to the group.(line 12)

Chapter 3

Experimental Results

3.1 Experimental setup

We studied the behavior of our constructive social choice algorithm on synthetic and realistic social choice tasks, both taken from [3][4]. Our experimental setup also follows closely the ones of [3] [4]. 20 groups of M = 20 users were randomly generated each using a hierarchical sampling procedure. Each group was split into C clusters, with M$\approx C$ users each. The users within a cluster are chosen to have similar preferences, simulating different sub-groups of users. For instance considering the scenario of buying a PC for a whole family or a group of users, there may be a cluster of power users who need more capable machines and a cluster of users who prefer energy efficient laptops. Having a cluster structure in this setting enables the transfer of preference information among similar users and also help to distribute the cognitive cost of query answering over multiple users for individual preference elicitation in MUSM. The implementation for social choice is noiseless, i.e. the group can not select the worst item from the presented queries at each iteration.

We implemented the *Constructive Social Choice algorithm* using Python, leveraging Gurobi 7.5.0 for solving the core MILP problem . All experiments were run on a 1,3 GHz Intel CPU with 5 cores and 8 GiB of RAM. For all experiments, one iteration corresponds to a single pairwise query. For all experiments the hyper parameter λ Was fixed to 0.5.

Dataset description

The first experiment is performed on the synthetic problem introduced in [3] for a social choice task over 20 groups each containing 20 users. In this setting the space of products $\mathcal{X}$ is taken to be the Cartesian product of r categorical attributes, each having r possible values. We use a one-hot encoding to represent products, for a total of r^2 0-1 variables. Here we focus on the $r = 4$ case with $r^2 = 16$ variables and $r^r = 256$ total products.

In the second experiment we consider a realistic PC configuration as in [3][4] for a social choice task over 20 groups each containing 20 users. The constructive recommendation is required to suggest a fully customized PC. A PC configuration is defined by 7 categorical attributes:

- Computer type (laptop, desktop, or tower)
- Manufacturer (8 choices)
- CPU model (37choices)
- Monitor size (8)
- RAM amount (10),
- storage size (10) and
- Price, *a linearly dependent numerical attribute defined as a linear combination of the other attributes.* Often the price of a PC is well approximated by the sum of the price of its components plus a bias due to branding.

The interactions between the attributes are expressed as Horn clauses, expressing statements like "manufacturer A does not sell CPUs of brand B"(i.e a certain manufacturer implies a set of possible CPUs), The PC dataset includes 16 Horn constraints leading to a search product space of about 700,000 distinct configurations.

3.1.1 Results

The true preference of the groups x^* is computed as :

$$x^* = \text{argmax}_x \sum_{u=1}^{M} \omega^*{}_u \langle w^*{}_u, x \rangle$$

The true user contributions ω^* are not directly observed, and must be estimated. The constructive social choice algorithm tracks an estimate of the user contributions at all times, iteratively improving it through group-wise interaction.

The constructive social choice algorithm first aggregates the preferences of individual users produced by MUSM then in every iteration it solves the objective function(obtaining configurations say x_1 and x_2 of high quality and diversity in order to present to the users),then these queries are presented to the group and the group as a whole gives feedback. The implementation is noiseless, i.e. the group can not select the worst item of the two.

The performance on a group is measured as the average regret over all N groups (*y-axis*). The regret(utility loss) is computed as :

$$\text{Regret}(x) = \sum_{u=1}^{M} \omega^*{}_u \langle w^*{}_u, x^* \rangle - \sum_{u=1}^{M} \omega^*{}_u \langle w^*{}_u, x \rangle$$

Synthetic setting: Figures 3.1 and 3.2 report the behavior of constructive social choice algorithm for 100 iterations over dense and sparse users sampled from standard normal and uniform distribution respectively.

It can be seen that the synthetic experiment on sparse users tends to favor earlier convergence over the first 2 to 3 iterations. A closer look at the convergence points is reported on figures 3.3 and 3.4.

For all experiments the computational cost is measured in terms of cumulative time. figures 3.5 and 3.6 report the cumulative time for for 100 iterations over dense and sparse users sampled from standard normal and uniform distribution respectively.

3.2 Limitation

As it is clearly seen from the experimental results the Constructive social choice algorithm works well for sparse users but fails to address dense users and it needs refinement to overcome this limitation over dense users.

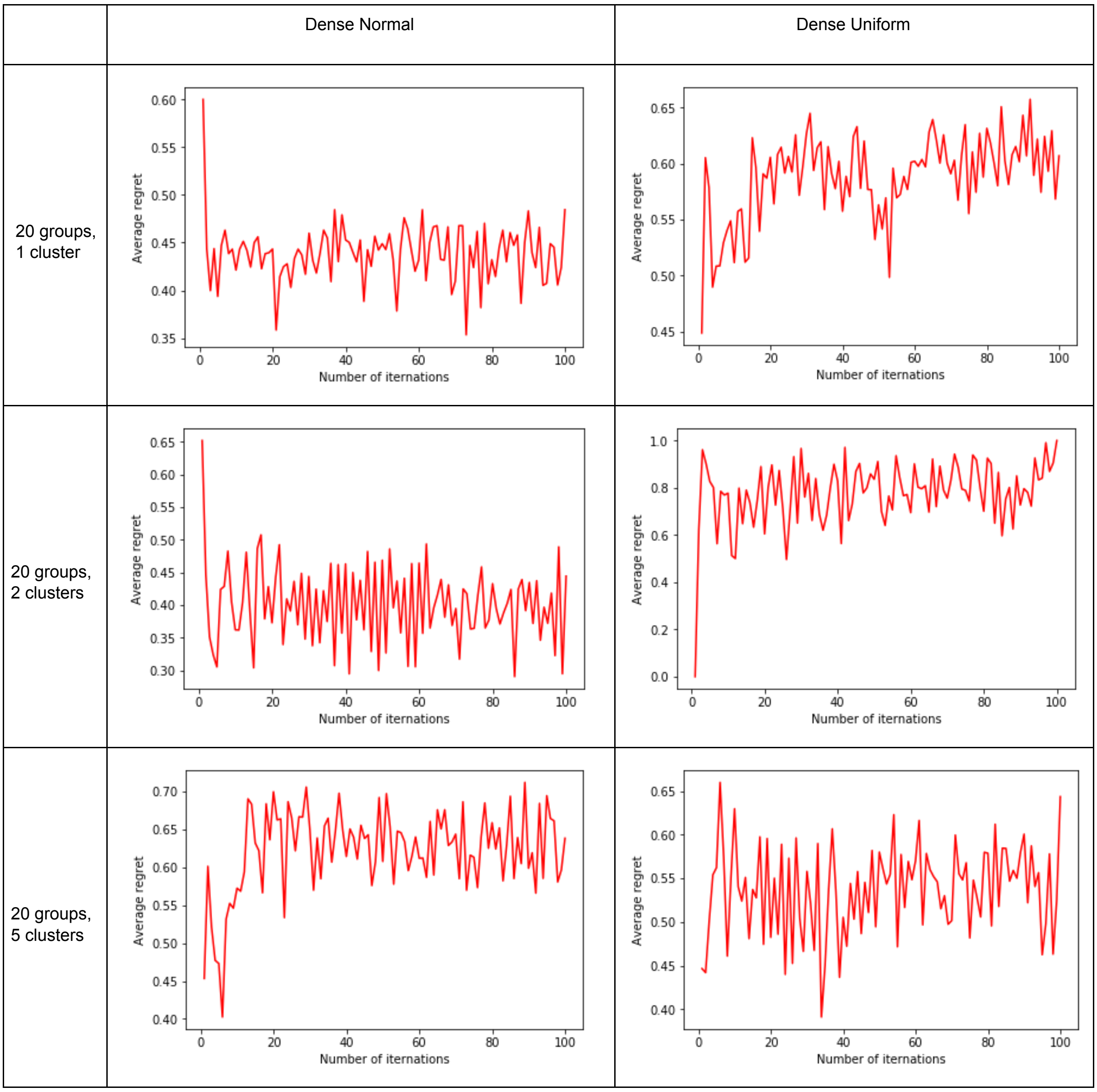


FIGURE 3.1: Synthetic experiment: average regret 20 Groups, dense users, 100 iterations

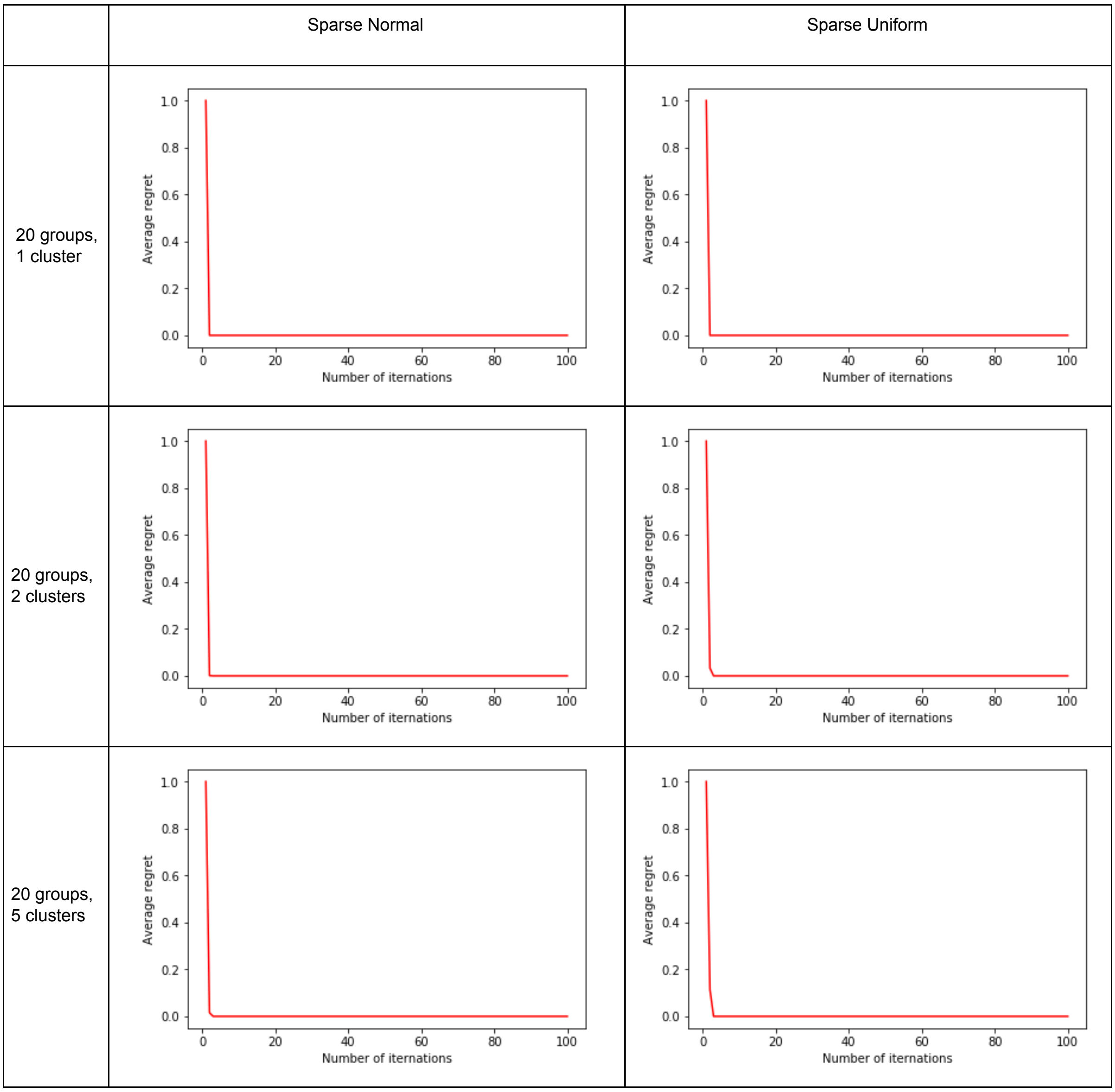


FIGURE 3.2: Synthetic experiment: average regret 20 Groups, sparse users, 100 iterations

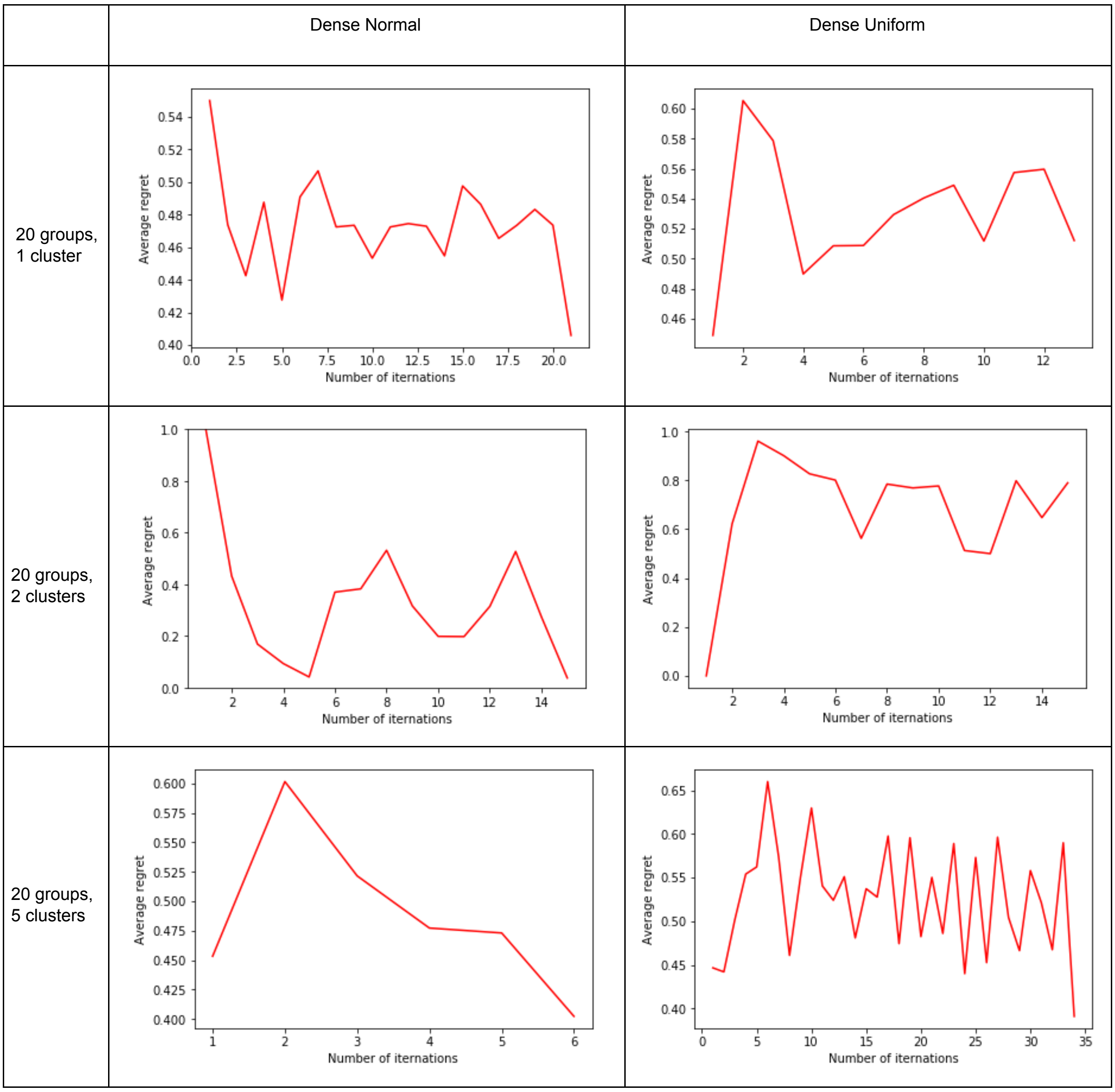


FIGURE 3.3: Synthetic experiment: average regret 20 Groups, dense users, convergence

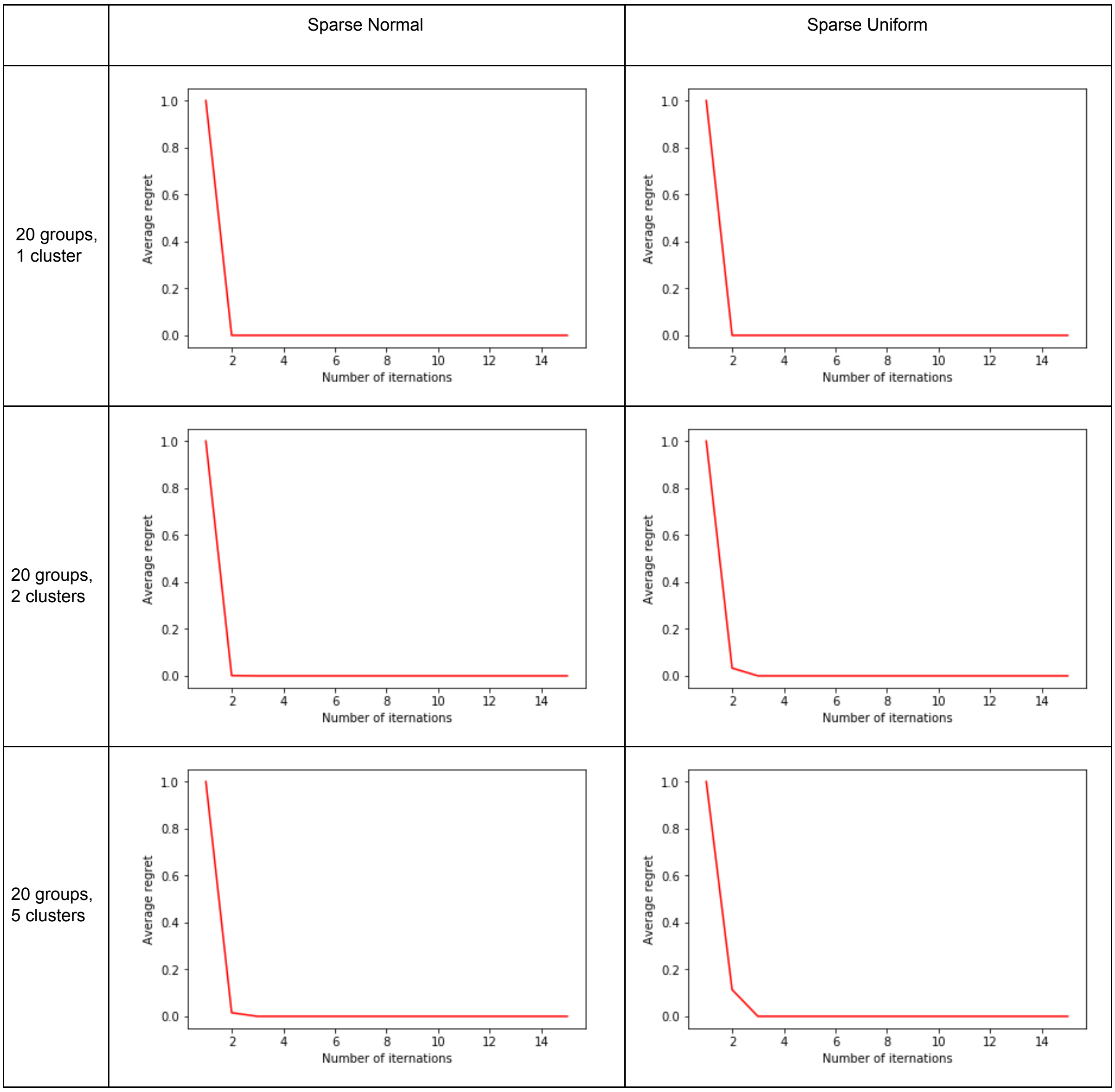


FIGURE 3.4: Synthetic experiment: average regret 20 Groups, sparse users convergence

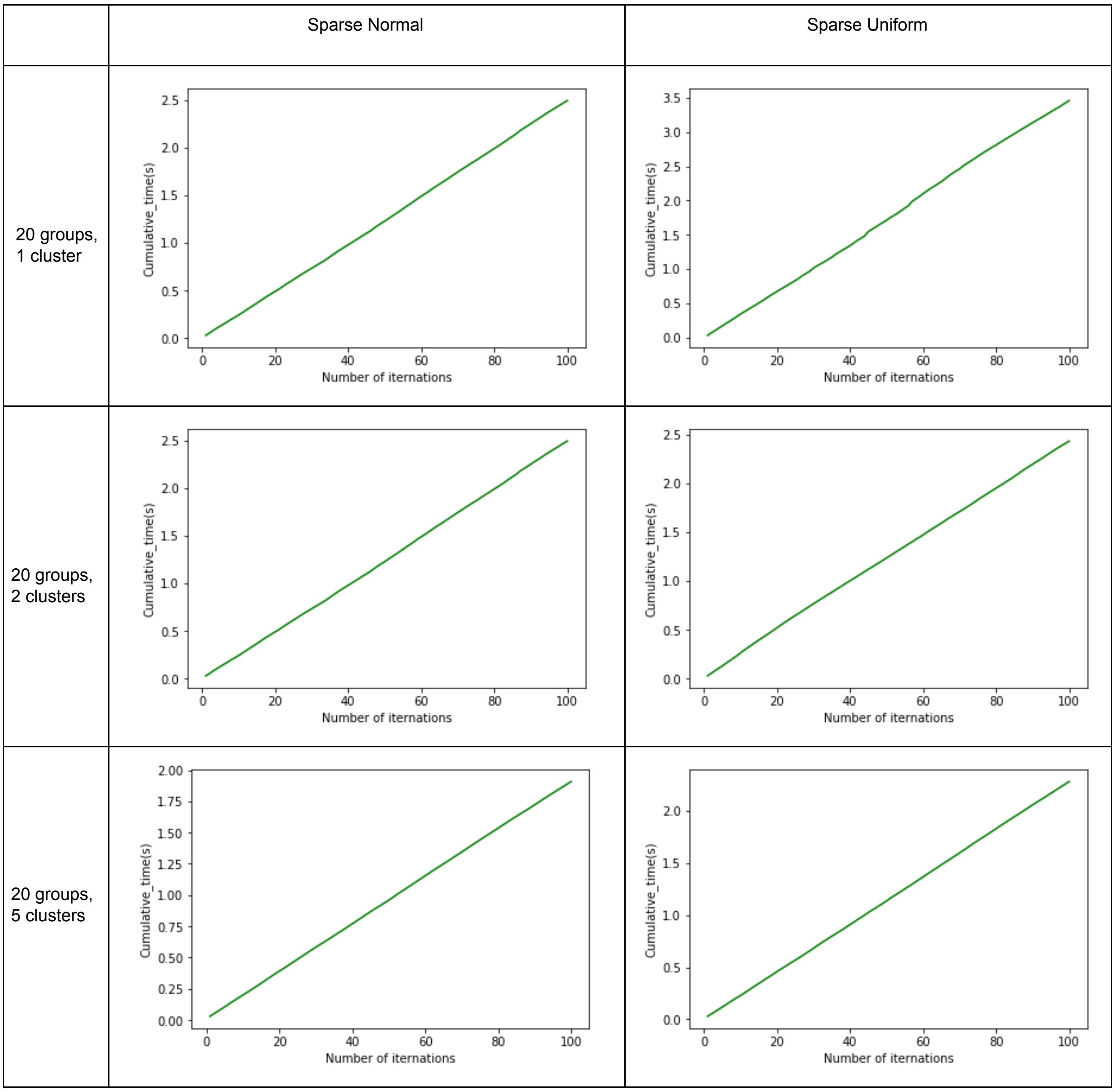


FIGURE 3.5: Synthetic Cumulative time 20 Groups, sparse users

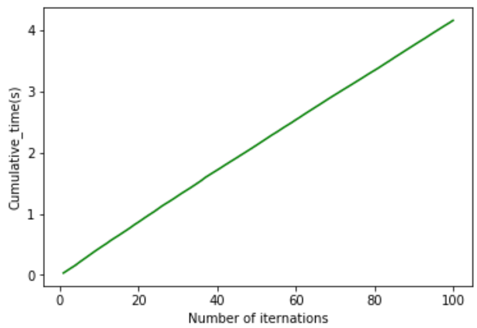
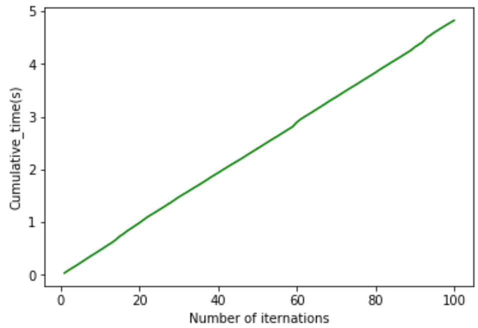
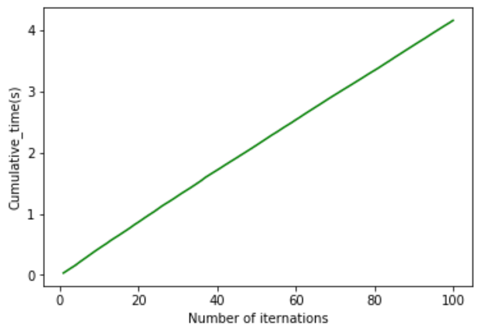
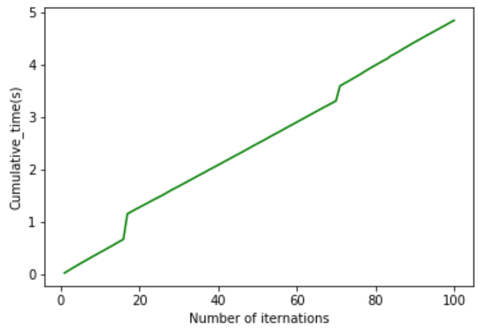
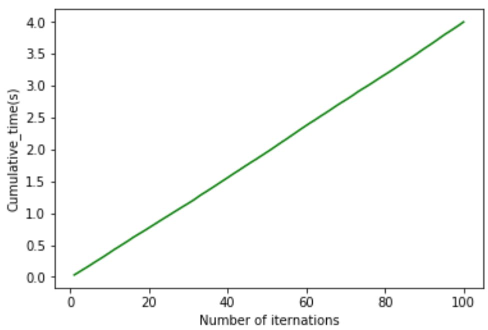
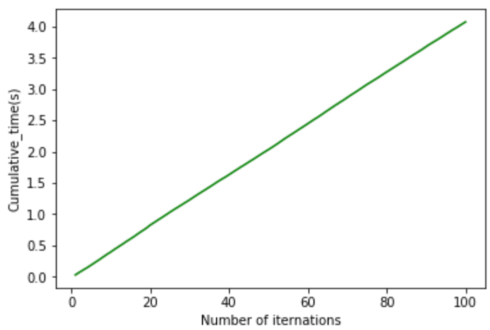
	Dense Normal	Dense Uniform
20 groups, 1 cluster		
20 groups, 2 clusters		
20 groups, 5 clusters		

FIGURE 3.6: Synthetic Cumulative time 20 Groups, dense users

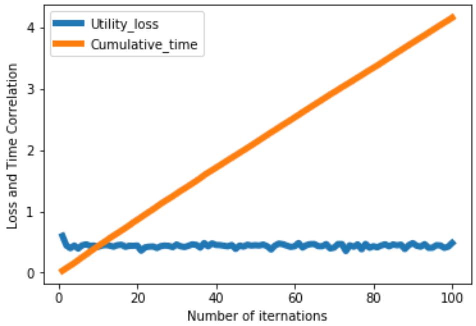
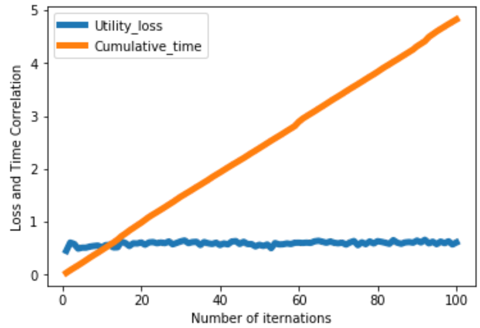
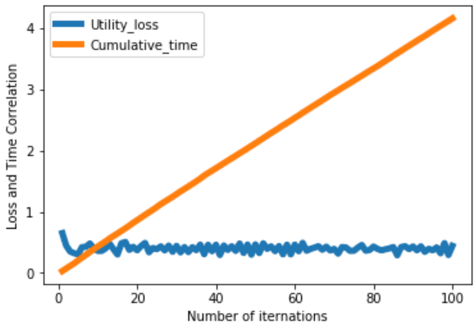
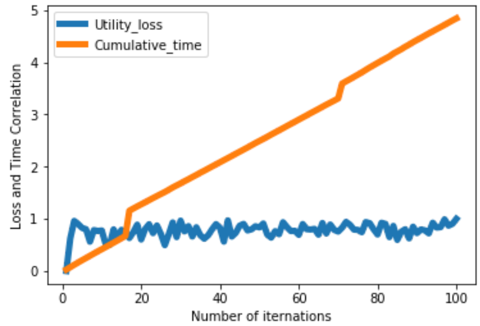
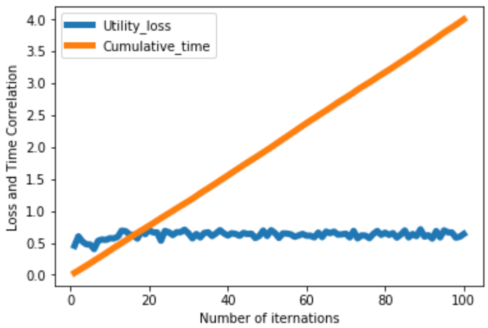
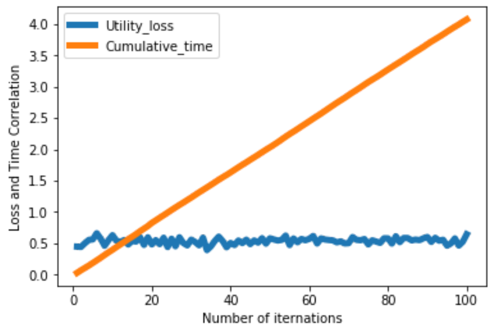
	Dense Normal	Dense Uniform
20 groups, 1 cluster		
20 groups, 2 clusters		
20 groups, 5 clusters		

FIGURE 3.7: Synthetic loss and time correlation 20 Groups, dense users

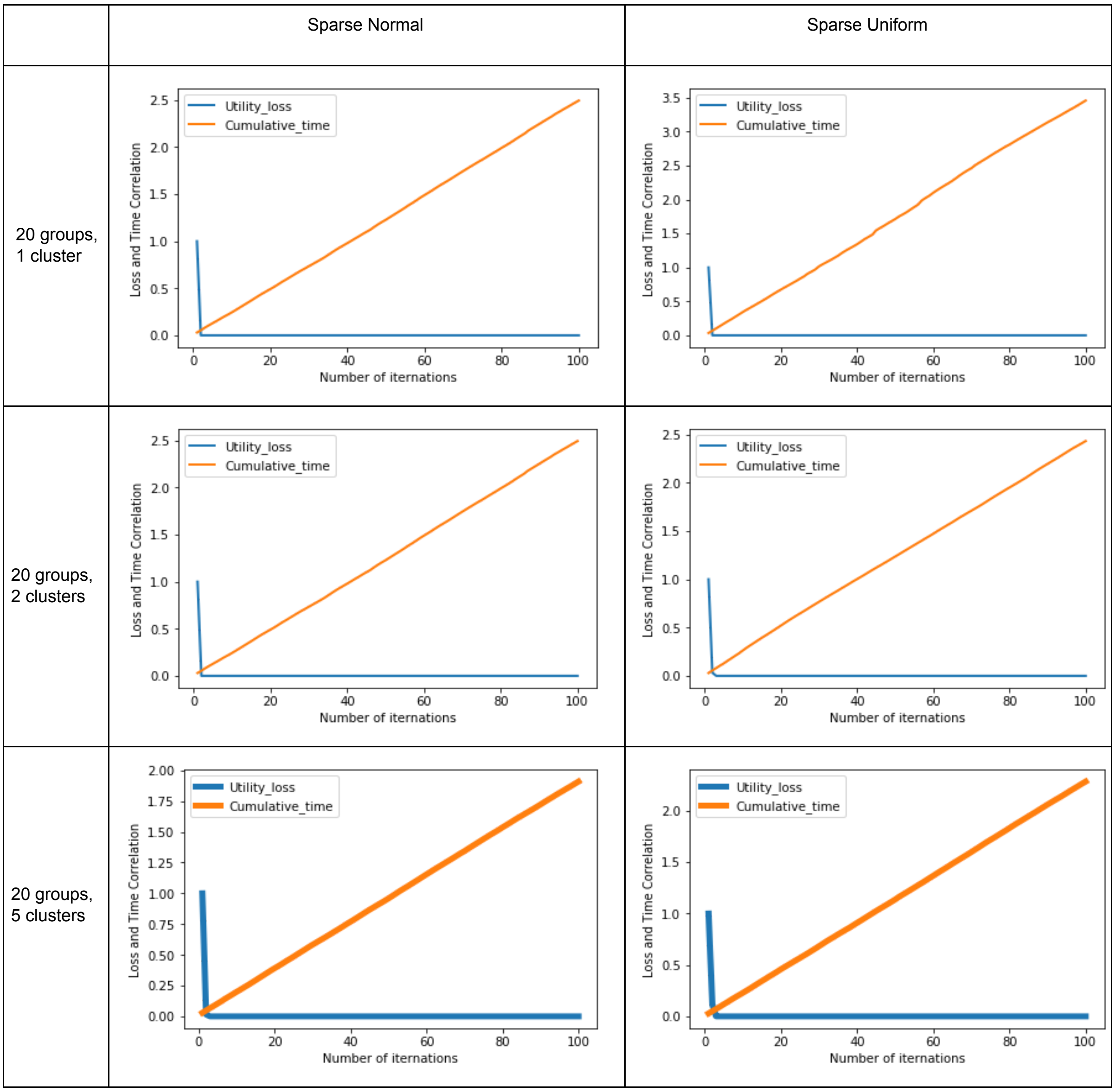


FIGURE 3.8: Synthetic loss and time correlation 20 Groups, sparse users

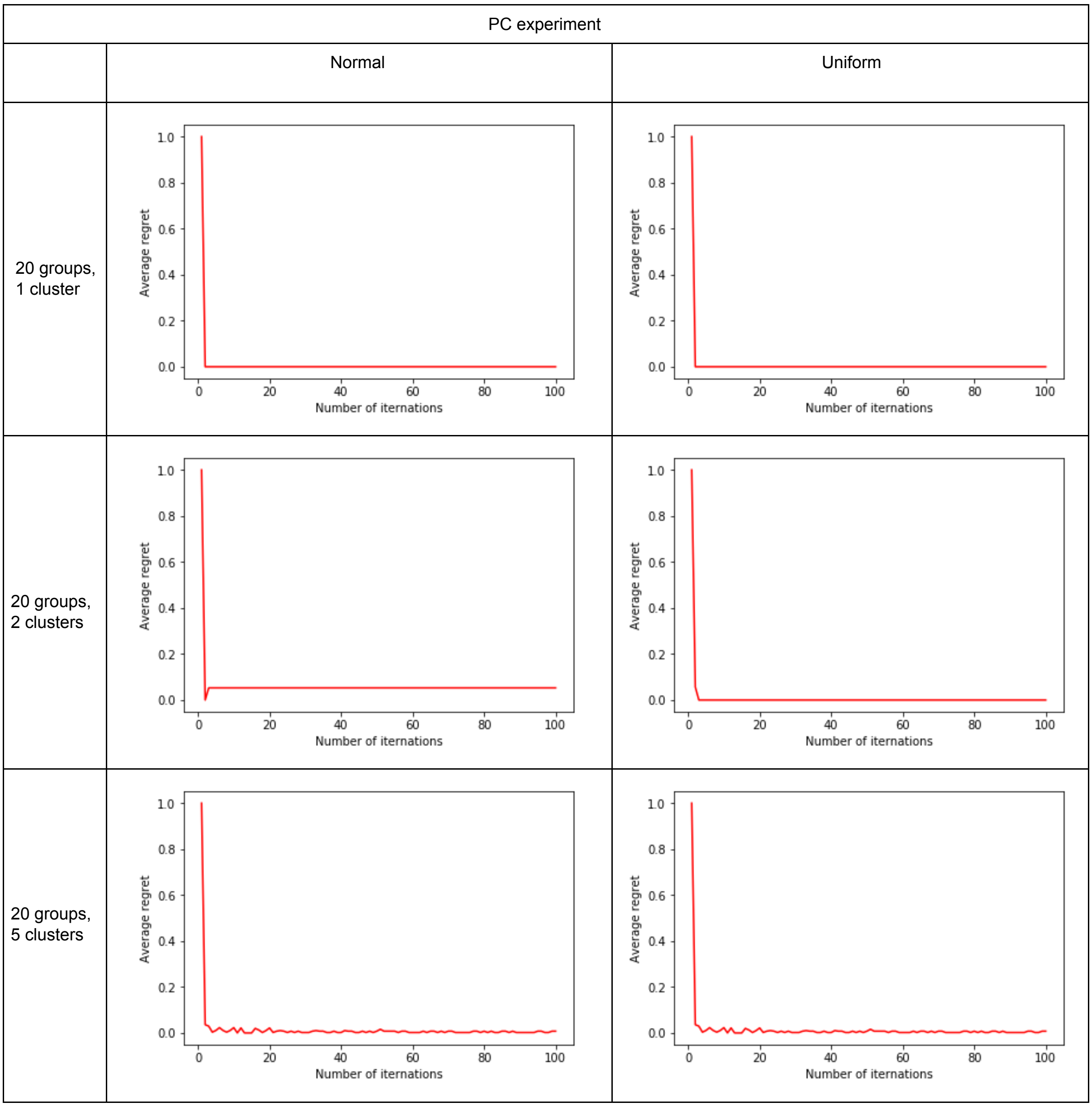


FIGURE 3.9: PC average regret 20 Groups, 100 iterations

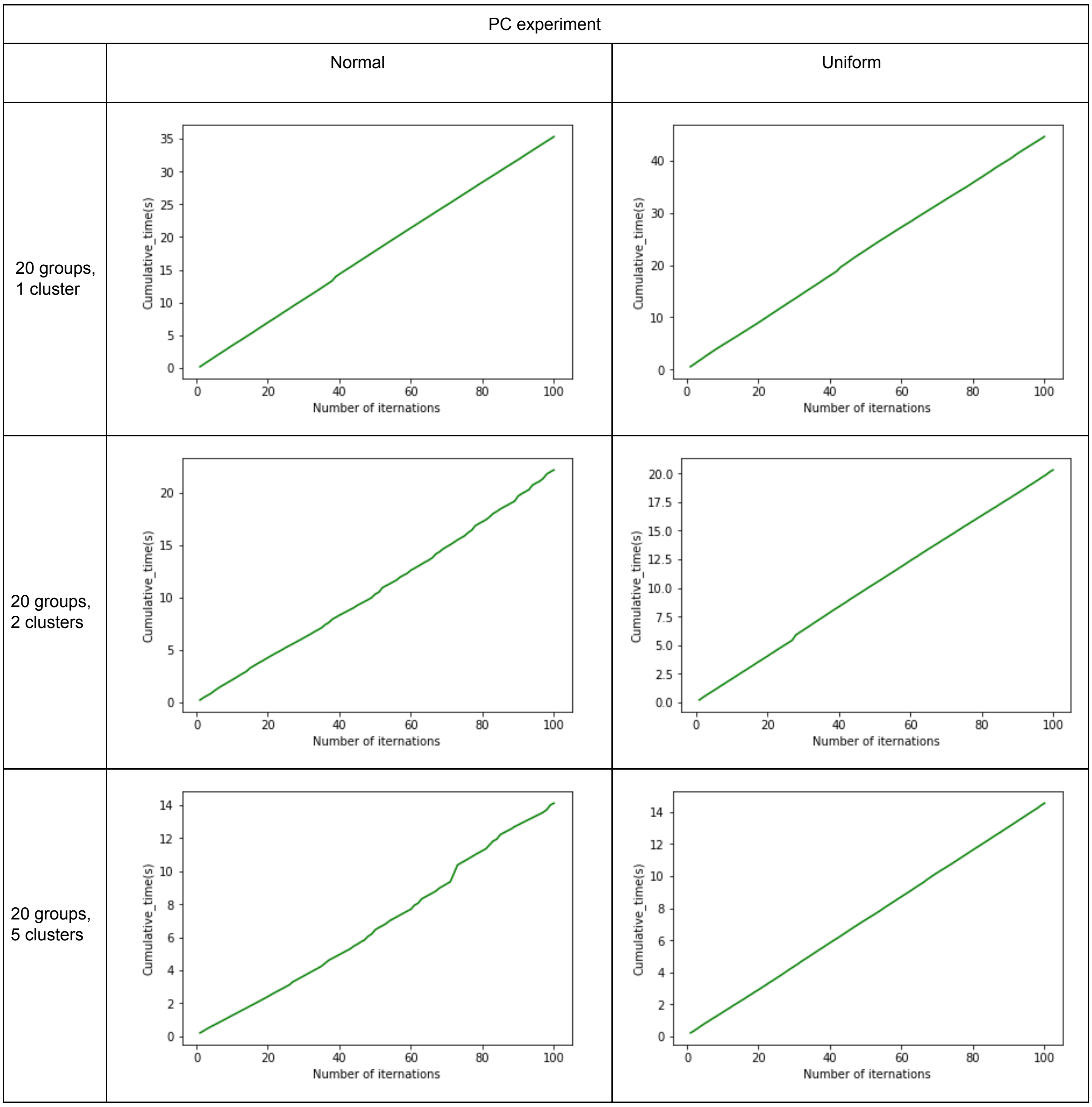


Figure 3.10: PC Cumulative time 20 Groups, 100 iterations

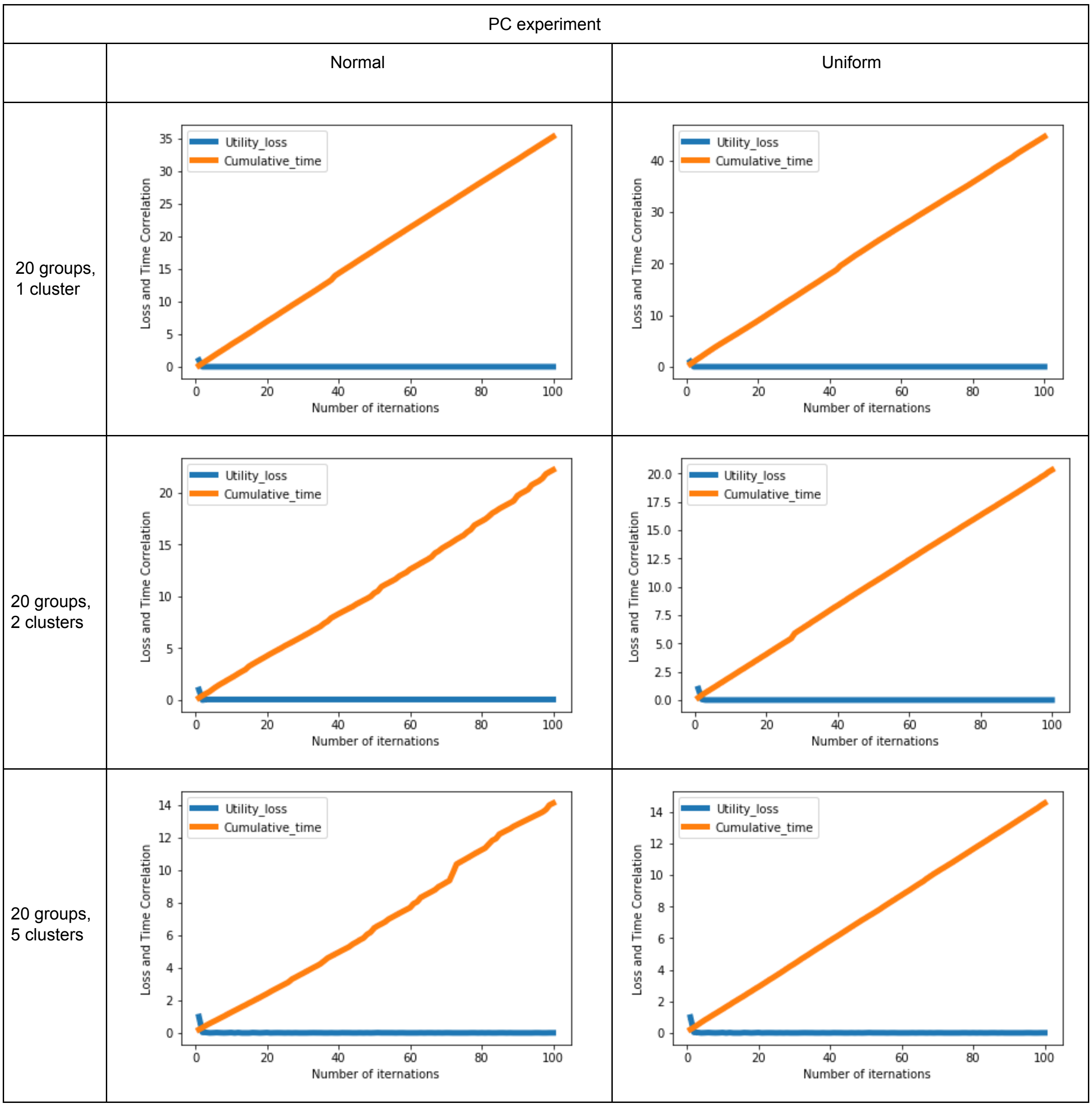


FIGURE 3.11: PC loss and time correlation 20 Groups, 100 iterations

Figure 3.12 shows a comparison between user groups having 1,2 and 5 clusters for the synthetic and the PC experiment. As mentioned in the above section a cluster with in a group represents users having similar preferences, simulating different sub-groups of users. For instance considering the scenario of buying a PC for a whole family or a group of users, there may be a cluster of power users who need more capable machines and a cluster of users who prefer energy efficient laptops. Having increasing number of cluster implies the existence of more diverge or may be conflicting interests with in a group which makes the group-wise decision objective more complex. Experimental results demonstrated that the constructive social choice algorithm converges earlier for decreasing number of clusters both for synthetic and PC experiments. Implying the more diverse the society the harder the social choice problem gets.

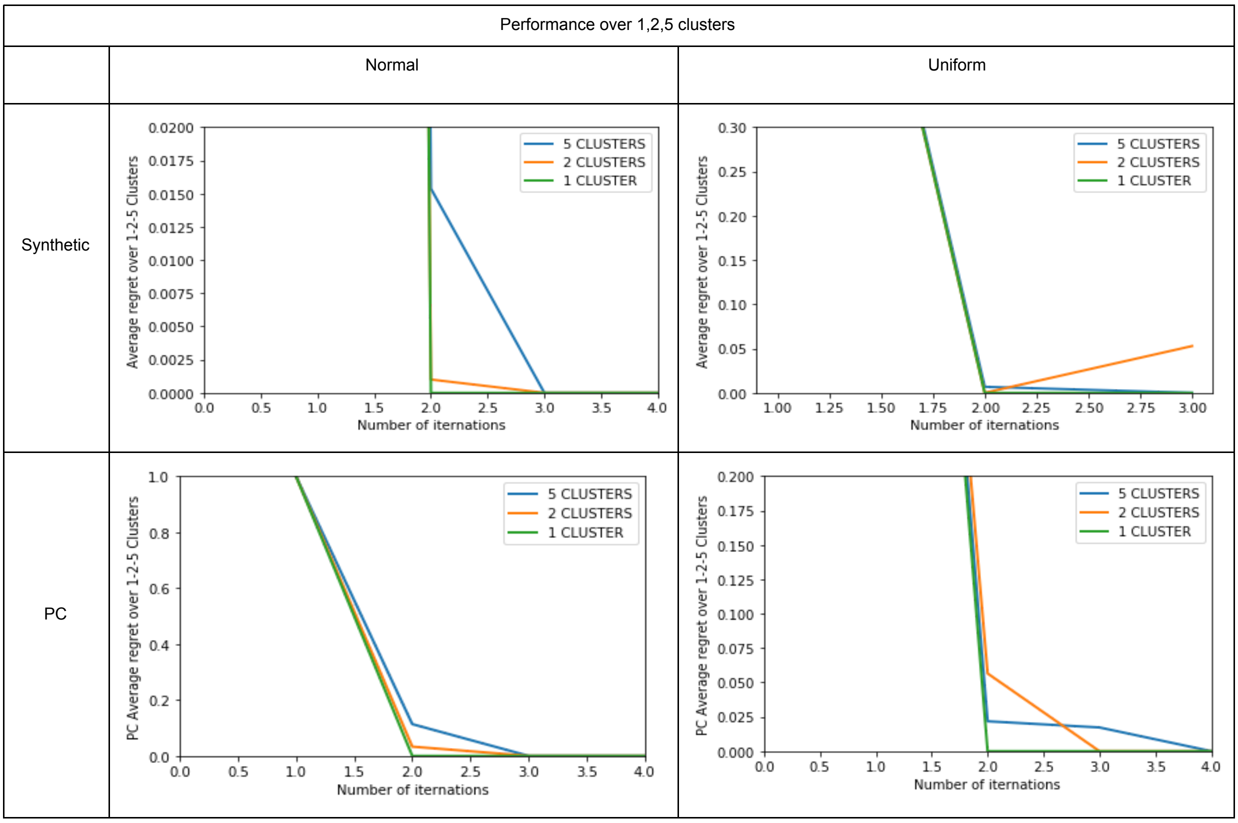


FIGURE 3.12: Performance over 1,2,5 clusters

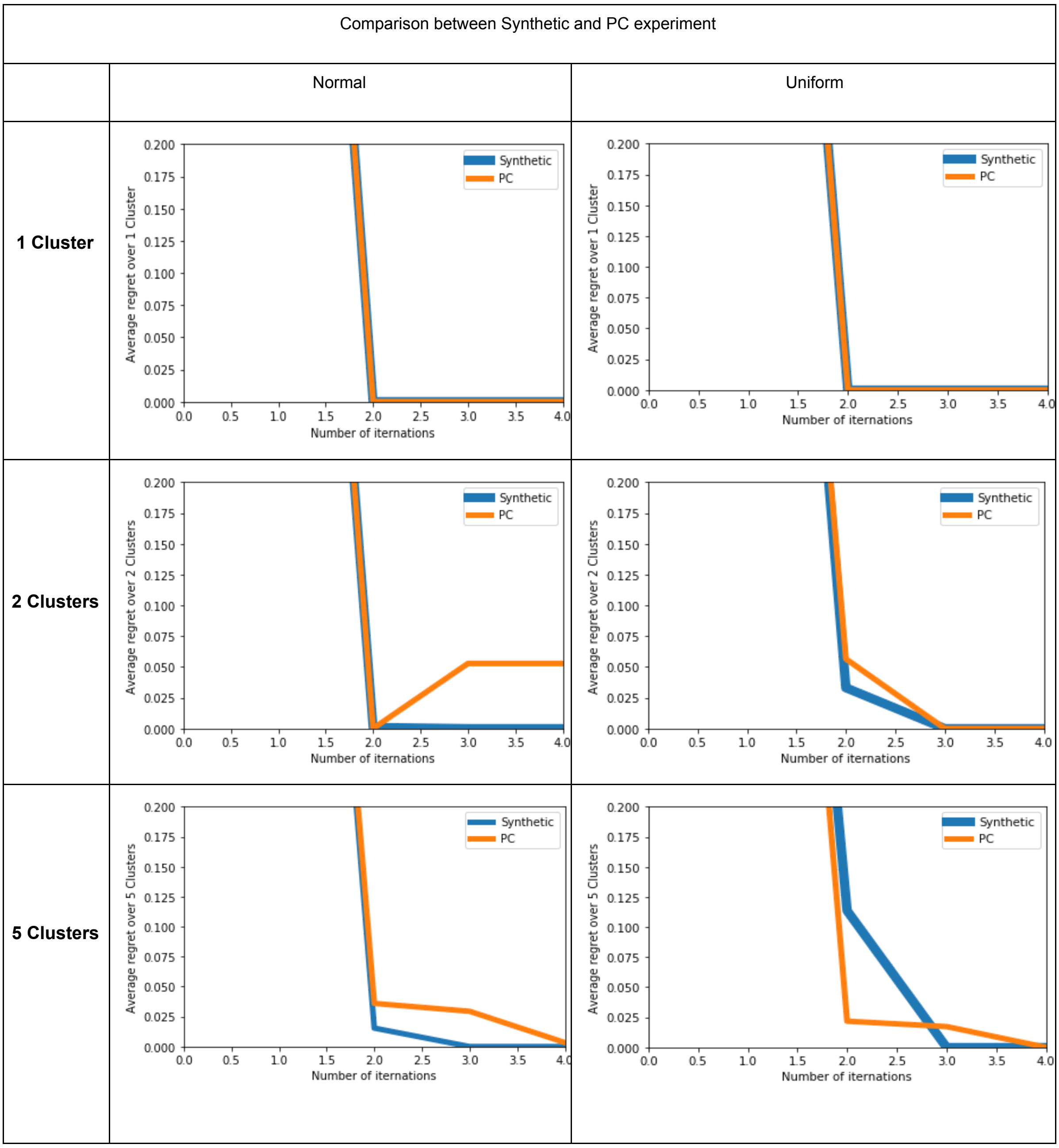


FIGURE 3.13: Comparison between Synthetic and PC experiment

It can also be seen that the constructive social choice algorithm yields almost indistinguishable results between synthetic and PC experiments on groups having 1 and 2 clusters of users and shows a slight variation with increasing number of clusters. from this observation it can be deduced that the constructive social choice algorithm can perform well especially for the more complex scenarios like the PC recommendation task.

Chapter 4

Conclusion and Future Work

4.1 Conclusion

We presented an approach to make a first step towards extending constructive preference elicitation to social choice domains. Classical social choice focuses neither on learning nor on recommendation but only on analyzing voting protocols, and no existing "social preference learner" is constructive and these approaches are incapable of solving the problem of social choice when the set of options are combinatorial(i.e the item to be recommended is a combination of sub parts like PC configurations, sets of shared exam schedules, home furnitures...etc making the items recommended to be generated from scratch instead of being chosen from among a set of available options). This requires robust constructive recommendation approach in order to make the best decision for the group or society.

We tackled this domain mainly because constructive learning in social choice is a natural extension and many problems can be measured in it. Since the problem of social choice has not yet been addressed from a constructive recommendation perspective we proposed a constructive social choice algorithm with set-wise max margin. A key idea is that, by asking informative questions and collecting group wise feedback to iteratively elicit societal preference and make the best recommendation to the society according to the learned feedback.

The main advantages of the constructive social elicitation method is ability to provide recommendations that maximize the consensus(utility) of a society in large configuration problems.

4.2 Future Work

As this is the first work in extending constructive preference elicitation to social choice domains we focused on leaning ω (the contribution of individual members) to the societal choice and as mentioned in the limitation section experimental results demonstrated that a test on dense users resulted a performance worsening scenario.
As a future work we propose

- An extension to learning both the preferences of individual users and their contributions (W and ω).
- Further refinement of the algorithm to address dense users.
- A deep experimental evaluation of more realistic and complex datasets like the PC dataset.

Bibliography

[1] Hensher, D. A., Louviere, J. J., Swait, J. D. . ”Conjoint preference elicitation methods in the broader context of Random Utility Theory preference elicitation methods” Institute of Transport Studies, Australian Key Centre in Transport Management, the University of Sydney and Monash University, 1999.

[2] List, Christian, Edward N. Zalta. ’Social Choice Theory.’ Metaphysics Research Lab, Stanford University.: Stanford Encyclopedia of Philosophy, Winter 2013.

[3] Teso, S., Passerini, A., Viappiani, P. Constructive preference elicitation by setwise max-margin learning. In: Proceedings of International Joint Conference on Artificial Intelligence IJCAI. pp. 20672073 (2016).

[4] Teso, S., Passerini, A., Viappiani, P. ’Constructive Preference Elicitation for Multiple Users with Setwise Max-margin.’ Algorithmic Decision Theory Lecture Notes in Computer Science, 2017.

[5] Marden, J. I. ’ Analyzing and modeling rank data.’ London: Chapman and Hall,(1995).

[6] Zimmermann, and Michael D. Intriligator. ’Probabilistic and Utility-theoretic Models in Social Choice.’ Review of Economic Studies, Oxford University Press. ideas.repec.org/a/oup/restud/v40y1973i4p553-560..html (Jan 1973).

[7] Boutilier, Craig.. ’Learning Mallows models with pairwise preferences.’ :Proceedings of the 28th International Conference on Machine Learning (ICML) Bellevue, WA, 2011.

[8] T. Lu and C. Boutilier. ’Robust approximation and incremental elicitation in voting protocols.’ IJCAI-11, Barcelona, 2011.

[9] YBlum, Avrim, et al. ’Preference Elicitation and Query Learning.’ Learning Theory and Kernel Machines Lecture Notes in Computer Science. : 2003.

[10] Baumol, William J., and Kenneth J. Arrow. ’Social Choice and Individual Values.’ Econometrica, vol. 20, no. 1, 1952, p. 110.

[11] Lahaie, Sebastien M., and David C. Parkes. ’Applying learning algorithms to preference elicitation.’ Proceedings of the 5th ACM conference on Electronic commerce - EC 04, 2004.

[12] Craig Boutilier and Tyler Lu. ’Probabilistic and utility-theoretic models in social choice: Challenges for learning, elicitation, and manipulation.’ In International Joint Conference on Artificial Intelligence IJCAI Workshop on Social Choice and Artificial Intelligence.

[13] Herrera-Viedma, E., et al. ’A Consensus Support System Model for Group Decision-Making Problems With Multigranular Linguistic Preference Relations.’ IEEE Transactions on Fuzzy Systems, vol. 13, no. 5, 2005, pp. 644–658.

[14] Boutilier, Craig. ’Preference Elicitation and Preference Learning in Social Choice.’ Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems Lecture Notes in Computer Science, 2011, pp. 1–1.

[15] Louviere, Jordan, et al. ’Conjoint Preference Elicitation Methods in the Broader Context of Random Utility Theory Preference Elicitation Methods.’ Conjoint Measurement, 2007, pp. 167–197.

[16] Dragone, Paolo. ’Constructive recommendation ’ In: Proceedings of the twenty sixth International Joint Conference on Artificial Intelligence IJCAI. (2017)