Towards Safety and Sustainability: Designing Local Recommendations for Post-pandemic World

Gourab K Patro
Indian Institute of Technology Kharagpur, India

Abhijnan Chakraborty
Max-Planck Institute for Software Systems, Germany

Ashmi Banerjee
Technical University of Munich, Germany

Niloy Ganguly
Indian Institute of Technology Kharagpur, India

ABSTRACT

The COVID-19 pandemic has made it paramount to maintain *social distance* to limit the viral transmission probability. At the same time, local businesses (e.g., restaurants, cafes, stores, malls) need to operate to ensure their economic sustainability. Considering the wide usage of local recommendation platforms like Google Local and Yelp by customers to choose local businesses, we propose to design local recommendation systems which can help in achieving both safety and sustainability goals. Our investigation of existing local recommendation systems shows that they can lead to overcrowding at some businesses compromising customer safety, and very low footfall at other places threatening their economic sustainability. On the other hand, naive ways of ensuring safety and sustainability can cause significant loss in recommendation utility for the customers. Thus, we formally express the problem as a multi-objective optimization problem and solve by innovatively mapping it to a bipartite matching problem with polynomial time solutions. Extensive experiments over multiple real-world datasets reveal the efficacy of our approach along with the three-way control over sustainability, safety, and utility goals.

CCS CONCEPTS

• **Information systems** → **Recommender systems**.

KEYWORDS

Safety, Sustainability, COVID-19, Social Distancing, Local Recommendation, Bipartite Matching, Yelp, Google Local

ACM Reference Format:
Gourab K Patro, Abhijnan Chakraborty, Ashmi Banerjee, and Niloy Ganguly. 2020. Towards Safety and Sustainability: Designing Local Recommendations for Post-pandemic World. In *Fourteenth ACM Conference on Recommender Systems (RecSys '20), September 22–26, 2020, Virtual Event, Brazil.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3383313.3412251

1 INTRODUCTION

With the proliferation of GPS-enabled smartphones, local recommendation platforms like Google Local (rendered on Google Maps), Yelp, Zomato, etc. have experienced massive growth in the last few years. For example, since 2011, the use of "near me" service on Google Local has grown by an astounding 3400% [39]. These platforms recommend nearby/local businesses (restaurants, cafes, stores, malls, etc.) to customers based on their physical locations and other inferred preferences, and in 2016, customers have visited around 1.5 billion businesses every month using these location-based services [39]. However, these regular customer-business physical interactions have been severely impacted due to the spread of highly contagious SARS-CoV-2 and the resultant COVID-19 pandemic. To limit the viral spread, many countries enforced complete/partial lockdowns for an extended period leading to the closure of several businesses, and even after reopening, strict adherence to social distancing guidelines is an absolute requirement to ensure safety of the customers. Considering the extensive use and influence of local recommendation platforms in attracting customers to local businesses, in this paper, we propose to design local recommendation systems which can help in achieving **safety for customers** as well as economic **sustainability for businesses** in the post-pandemic world.

Traditionally, these platforms have used different data-driven models [12, 21, 24, 28, 34, 44] to estimate relevance of local businesses to individual customers, and then recommended k most relevant results to them. By gathering data from Google Local and Yelp, we show that such pre-COVID recommendation practices can cause a high inequality in the exposure (visibility) of local businesses, where a few businesses can end up receiving a large fraction of total exposure while the remaining businesses receive a very low exposure. This could, on one hand, lead to overcrowding at some businesses, compromising customer safety. On the other hand, it could result in a very low footfall at other businesses, questioning their sustainability in the ongoing scenario (detailed in §3). A simple answer to these concerns would be to find a way which can reduce inequality in business exposures. However, using naive methods to reduce exposure inequality (e.g., poorest-k: recommending k least exposed businesses to customers) may result in a huge loss in customer utility (detailed in §3), thereby rendering the platform inefficient for customers. Therefore, our focus on safety and sustainability need to go hand-in-hand with customer utility.

We formally define the desired properties for sustainability, safety, and utility in section 3.1. For sustainability, we propose to use a minimum exposure guarantee for every business, and for safety, we propose to keep the exposure of a business below a certain maximum limit which is proportional to its safe capacity. As we observe in the case of poorest-k, there is a clear tradeoff between utility and sustainability/safety, and simultaneously satisfying all

RecSys '20, September 22–26, 2020, Virtual Event, Brazil

ACM ISBN 978-1-4503-7583-2/20/09…$15.00
https://doi.org/10.1145/3383313.3412251

the constraints is not possible. Thus, we relax the constraints into the following three optimizable objectives: (i) minimize exposure deficit for business sustainability (ii) minimize exposure surplus for customer safety, (iii) minimize utility loss (more details in section 4.1). We combine these objectives to formulate a multi-objective optimization problem and adopt a novel way to map it to a bipartite matching problem with polynomial time solutions (detailed in section 4.3). Next, we test our local recommendation mechanism on multiple datasets and evaluate the results with several metrics (section 5.1) capturing various aspects of local recommendations. Extensive evaluations reveal the efficacies of our approach along with the three-way control over sustainability, safety, and utility goals.

In summary, we make the following contributions in the paper: (i) we consider very timely and much needed notions of business sustainability and customer safety in local recommendation systems, and formally define these notions along with customer utility requirement (section 3.1)— to our knowledge we are the first to do so; (ii) we incorporate these goals into a multi-objective optimization framework and solve by innovatively mapping it to a bipartite matching problem with polynomial time solutions (§4); (iii) we empirically test and evaluate it on multiple real-world datasets gathered from platforms like Foursquare, Google Local and Yelp to show the effectiveness of our solution (§5). We believe that such local recommendation systems designed with sustainability and safety goals would be a very effective complement to other location-based services like contact tracing applications [27, 37] aggresively recommended by various governmental and non-governmental organizations to abate inconveniences during these unprecedented times.

2 PRELIMINARIES

2.1 Datasets Gathered

In this paper, we gather the recommendations patterns of local businesses in New York City **(NYC)** and San Francisco **(SF)** on two platforms: Google Local and Yelp, by using a publicly available dataset on customer checkins in FourSquare. Note that these recommendations were gathered during pre-COVID times in 2019.

Customer Locations (Foursquare Check-in Data): To get an estimate of the locations from where customers are accessing location-based platforms and looking for nearby businesses, we use a publicly available Foursquare check-in data from NYC and SF [41, 42]. The dataset contains 227, 428 and 572, 338 check-ins posted by customers at different restaurants around NYC and SF respectively, along with their geographic coordinates (i.e., latitude & longitude). We treat these check-in coordinates as customer locations (from where they log into platforms like *Yelp*, *Google Local*), and attempt to collect the recommendations provided by the platforms.[1]

[1]Collecting the recommendation results for every check-in location is not possible due to API limits. Hence, we cluster the check-in locations of each city into $K = 1000$ clusters using K-Means clustering [2], where the average cluster diameters came out to be 498 meter for NYC and 199 meter for SF. While collecting platforms' recommendations, we consider the centroids of these clusters as the customer locations. Number of check-in locations in each cluster serves as the popularity metric of the corresponding location (cluster centroid).

Yelp Dataset: *Yelp.com* platform for local restaurants is powered by a crowd-sourced review forum. For the customer locations described earlier, we collect nearby restaurants recommendations using Yelp Fusion API (`yelp.com/fusion`). The ranked lists contain the name and geographic coordinates of the restaurants, their distance from customer location, average rating, etc. In total, we have the data on 5702 and 3587 restaurants in NYC and SF respectively. Henceforth, we refer to the Yelp datasets of New York and San Francisco as **YP-NYC** and **YP-SF** respectively.

Google Local Dataset: Using a process similar to *Yelp*, we gather recommendation results from *Google Local* (a nearby recommendation service rendered on Google Maps) for the customer locations in NYC and SF using Google Places API [17]. In total, we have data on 2087 and 1478 restaurants in NYC and SF respectively. We refer to the datasets of New York and San Francisco as **GL-NYC** and **GL-SF** respectively.

2.2 Notations

In this paper, U and P refer to the set of customers and local businesses listed on the platform while u, p are instances of customer and business respectively, and $|U| = m$, $|P| = n$. We use $\mathcal{I}$ ($<u^i, l^i> \in \mathcal{I}$) for the sequence of login instances, where any i^{th} login instance is a tuple $<u^i, l^i>$ representing the login by customer u^i from location l^i. Each login instance is a single customer session on the platform. For login instance i, let $\mathcal{R}^i$ represent the ranking/permutation of the businesses in the recommendation, $\mathcal{R}^i_p$ represent the position/rank of business p in $\mathcal{R}^i$, and $\mathcal{R}^i[j]$ represent the j^{th} ranked business in $\mathcal{R}^i$. We use **index notations**: i for login instances, and j for rank/positions in recommedations. Superscripts represent the corresponding login instances; subscripts always represent the corresponding businesses.

2.3 Customer Utility of Ranked Recommendation

As we consider ranked recommendations, there are two important factors which drive customer utility: (i) relevance of businesses, and (ii) attention distribution over different ranks. We define them first before defining customer utility.

Relevance of a Business: The relevance of a business to a certain customer from a certain location, represents how likely the customer is going to visit the business and get a satisfying experience. Platforms often employ various data-driven algorithms (e.g., collaborative filtering [24, 44], content-based filtering [28, 34], learning-to-rank [23] etc.) to estimate the relevance scores. Let V be the relevance function, and $V^i(p) = V^i_p$ represent the relevance of business p to i^{th} login instance. Moreover, V^i_p can be thought of as the amount of utility u^i gains if she is recommended p when she logs in from location l^i. Note that we assume all the relevance scores to be non-negative.

Attention Distribution over Ranks: Prior works have shown that customers pay varied attention to differently ranked items and the overall attention distribution follows a drop-off after each rank/position [1, 11, 20]. In this paper, we consider standard logarithmic drop-off [20] for attention weights. The attention weights are given as below.

$$a(j) = \begin{cases} \dfrac{1}{J \cdot \log_2(j+1)} & 1 \leq j \leq k \\ 0 & j > k \end{cases} \quad (1)$$

Here $J = \sum_{j'=1}^{k} \dfrac{1}{\log_2(j'+1)}$ is a normalization parameter, and $a(j)$ is the normalized amount of attention received by a business in position j. We consider non-zero attention weights only for positions 1 till k based on the assumption that low-ranked subjects are hardly inspected. Note that every customer is assumed to follow the same attention distribution.

Customer Utility: For a specific login instance i, customer utility or recommendation utility will be high if highly relevant businesses are shown in top ranks. Thus, we use the normalized discounted cumulative gain: $NDCG@k$ metric [20] for customer utility (ϕ). The utility of $\mathcal{R}^i$ for login instance i can be defined as below.

$$\phi(\mathcal{R}^i, i) = NDCG@k(\mathcal{R}^i, i) = \frac{DCG@k(\mathcal{R}^i, i)}{IDCG@k(i)} \quad (2a)$$

where

$$DCG@k(\mathcal{R}^i, i) = \sum_{j=1}^{k} V^i\big(\mathcal{R}^i[j]\big) \times a(j) \quad (2b)$$

$$IDCG@k(i) = \max_{\mathcal{R}^i} \Big\{DCG@k(\mathcal{R}^i, i)\Big\} = DCG@k\big(\mathcal{R}^{i*}, i\big) \quad (2c)$$

and $\mathcal{R}^{i*}$ is the permutation of businesses in descending order of their relevances V_p^i.

When the businesses are sorted in descending order of their relevance, the DCG will be the highest – i.e., the $DCG@k$ of $\mathcal{R}^{i*}$ will be the highest. Thus, the range of customer utility $NDCG@k$ remains $[0, 1]$.

2.4 Relevance Estimation through Customer Survey

Current location-based platforms gather different kinds of data on the customer behaviours and preferences, on which various data-driven algorithms can be run to estimate relevance scores. In absence of such data, we run a survey with 140 Amazon Mechanical Turk (`mturk.com`) workers to gather data on how customers view different features like rating, distance, cuisine, etc. while selecting a nearby restaurant to visit. Each survey participant was asked to rank top 5 restaurants from a list containing 10 different restaurants with different ratings, distance and cuisine types; such top 5 ranking was asked from each participant for 7 times using 7 different lists. We use this data for personalized relevance estimation. For the purpose of this paper, we consider customer preferences based on distance and ratings of restaurants while estimating relevance scores. We employ a simple linear model (given by following equation) for point-wise learning-to-rank [22] to estimate personalized relevance scores.

$$V_p = w_0 + w_1 \times \left(\frac{n+1-\mathcal{R}_p^{\text{rating}^*}}{n}\right) + w_2 \times \left(\frac{n+1-\mathcal{R}_p^{\text{proximity}^*}}{n}\right) \quad (3)$$

where $\mathcal{R}^{\text{rating}^*}$ and $\mathcal{R}^{\text{proximity}^*}$ are the permutations of restaurants in descending order of ratings and descending order of proximity (or ascending order of distance) correspondingly. Here $n = 10$ as we provide 10 different options in each set. We learn $W = (w_0, w_1, w_2)$ for each of the participants separately. On testing, we find our model to achieve a mean $NDCG@5$ of 0.951. We use these learned customer models (W) for relevance estimation further in the paper. However, the platforms can easily replace above method with their own state-of-the-art relevance estimation method requiring no change in further formulation, as the relevance scores work just as inputs to our recommendation mechanism.

2.5 Business Exposure

The exposure of a business is the total amount of attention it gains from the customers; it explicitly depends on the business's position in each of the recommendations over time. The total exposure gained by a business p till t^{th} login instance can be given as below.

$$E_p^t = \sum_{i \in \mathcal{I}[1:t]} a\big(\mathcal{R}_p^i\big) \quad (4)$$

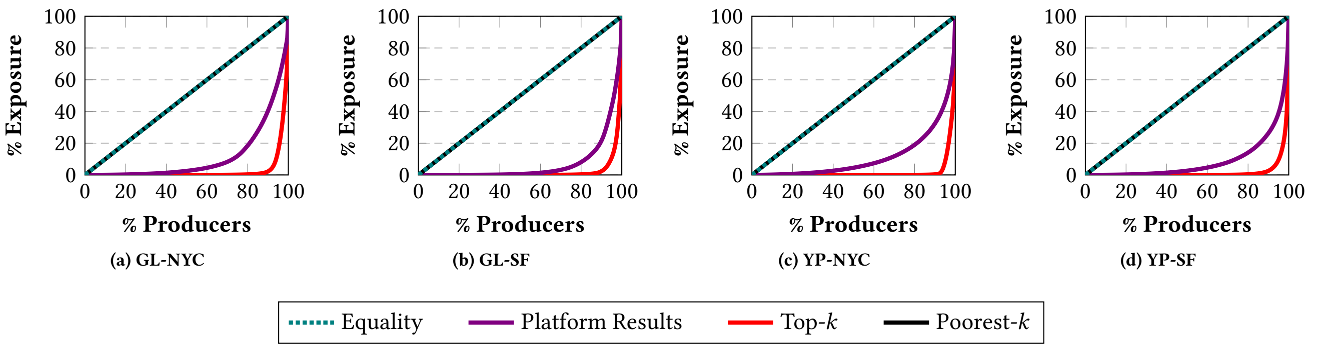
Note that each login instance is assumed to gather a total of 1 attention and thus 1 exposure.

3 MOTIVATION

Next, we simulate a regular local recommendation service. We consider the customers to be appearing in a random order, and each of the login instances is assumed to follow a randomly chosen customer model (W) learned in section 2.4. For each instance, first we consider two recommendation schemes: (i) the platform's recommendation (explained in section 2.1) and (ii) top-k (k most relevant businesses based on customer model W) recommendation. We record the total exposure received by each of the businesses over all the customer instances in each scheme.

Exposure Inequality in Conventional Top-k Recommendation: We plot Lorenz curves for business exposures in figures-1a, 1b, 1c and 1d. In these curves, the cumulative fraction of total exposure is plotted against the cumulative fraction of the number of corresponding businesses (ranked in increasing order of their exposures). The extent to which the curve goes below a straight diagonal line (or an equality mark indicated by light-green broken lines) indicates the degree of inequality in the exposure distribution. In all the cases, we see the platform's recommendations and also the top-k recommendations using our customer modeling show huge exposure inequalities; this leads to a small fraction of businesses getting most of customer attention while a majority of the businesses receive very small amount of exposure. Exposure determines economic opportunities in such platforms. Thus, low exposure could very well mean less footfall and less business. This raises a question on the sustainability of the businesses. On the other hand, very high exposure at some of the businesses could lead to huge footfall and overcrowding at those places, increasing the risk of viral spread. Therefore, we need a recommendation mechanism which could answer both sustainability and safety concerns. A simple answer to sustainability and safety concerns would be to reduce inequality in business exposures; and a naive approach for that would be poorest-k recommendation: recommend the k least exposed businesses at each instance.

Limitations of Poorest-k Recommendation: In figures-1a,1b,1c,1d, we see the Lorenz curves for poorest-k is very close to the equality mark. This proves that poorest-k could solve the sustainability and



(a) GL-NYC (b) GL-SF (c) YP-NYC (d) YP-SF

Figure 1: Lorenz curves show high inequality in exposure of businesses.

safety issue by distributing the total exposure almost equally; however, poorest-k reduces overall customer utility by 60-70% (observed in all the datasets) which could result in customers not liking the recommendations and a reduced usage of the platform; customers not following the recommendations could not only defy our goal of sustainability and safety, but also question the very survival of location-based platforms. Therefore, there is a need of recommendation mechanism which can address concerns for sustainability, safety along with overall customer utility.

3.1 Desiderata for Sustainability, Safety, and Utility

Here we formally define the necessary properties for sustainability, safety, and utility in recommendation.

Sustainability of businesses: For sustainability, we propose to ensure a minimum exposure guarantee for every business. This is comparable to the fairness of minimum wage guarantee [14, 18, 31]), which has been found to decrease income inequality [13, 25]. We also hope to reduce exposure inequality here. Formally, we write the criterion as below.

$$E_p \geq \overline{E}, \forall p \in P \tag{5}$$

where $\overline{E}$ is the minimum exposure guarantee. For online scenario, we can define the exposure guarantee as a moving guarantee: $\overline{E}^i = \beta i$, where $\overline{E}^i$ is the amount of exposure to be guaranteed to every business by i^{th} customer login instance; so β becomes the fraction of total exposure to be guaranteed as total exposure till i^{th} instance is i. As there are n businesses in total, we can limit setting β to $\beta_{\max} = \frac{1}{n}$.

Safety of Customers: For customer safety concerns, we propose to have a maximum limit on business exposure[2] and keep the maximum limit proportional to the safe capacity of the business.[3]

[2]Safety measurements can also make use of crowdsourced hygiene-standards, proximity of nearby infected clusters using existing contact-tracing apps, etc. However, in this paper we focus on the business exposure or expected crowd as an indicator for customer safety.

[3]From the floor area statistics of restaurants in the USA (https://www.statista.com/statistics/587130/average-floor-space-qrs-us/), we find that it varies in between 3000-4500 square feet. We assume 50% of that space comes available for seating at every business and take 2 metres as the prescribed safe social distance. Using maximal occupancy tool (https://covid19.mpi-sws.org/capacity_estimation/), we arrive at safe capacity range of 26-40 persons. In absence of true maximal capacity, we randomly sample an integer valued safe capacity from [26, 40] for each business.

This can reduce the chances of overcrowding at businesses thereby aiding in social distancing and enhancing customer safety. We can formally define this as below.

$$E_p \leq \zeta_p, \forall p \in P \tag{6}$$

where ζ_p is the maximum exposure for p and $\zeta_p \propto \mathit{capacity}(p)$. $\zeta_p = \frac{(\text{capacity of } p)\times(\text{total exposure})}{\text{total capacity of all businesses}}$.

Customer Utility: From the perspective of customers, the recommendations need to be relevant to them. Formally:

$$NDCG@k(\mathcal{R}^i, i) = 1, \forall i \in L \tag{7}$$

4 RECOMMENDATIONS TO INDUCE SUSTAINABILITY AND SAFETY

We design an online recommendation mechanism with sustainability, safety, and utility goals, i.e., at each customer login instance i, we need to find a ranked recommendation $\mathcal{R}^i$ with the goals in mind. As seen in the case of poorest-k (§3), there is a clear trade-off between utility and safety/sustainability, so all the three goals may not be achieved together. Thus we plan to relax the objectives and then combine them to form a joint optimization problem.

4.1 Relaxed Objectives

Here, we formally define the relaxed objectives for sustainability, safety, and utility.

Minimize Exposure Deficit: Instead of a hard constraint on exposure guarantee (eq-5 for sustainability), we propose a relaxed objective using exposure deficit. We define exposure deficit of a business as the relative difference between her exposure and minimum exposure guarantee if she is lagging behind, and 0 otherwise: i.e., $\max\{0, \frac{\overline{E}-E_p}{\overline{E}}\}$. We would like to minimize the exposure deficit of all the businesses. Following a utilitarian approach, we formulate a *min-sum* objective (minimize the sum of exposure deficits of all the businesses). For i^{th} login instance, it is as given below.

$$\underset{\mathcal{R}^i}{\arg\min} \sum_{p \in P} \max\left\{0, \frac{\beta i - \left(E_p^{i-1} + a(\mathcal{R}_p^i)\right)}{\beta i}\right\} \tag{8}$$

Minimize Exposure Surplus: Instead of a hard constraint on exposure based on safe capacity (eq-6 for safety), we propose a relaxed objective using exposure surplus. We define exposure surplus of a business as the relative difference between her exposure and her maximum exposure limit if she has a surplus, and 0 otherwise: i.e., $\max\{0, \frac{E_p-\zeta_p}{\zeta_p}\}$. Here also we follow a utilitarian approach, and formulate a *min-sum* objective for i^{th} login instance.

$$\underset{\mathcal{R}^i}{\arg\min} \sum_{p\in P} \max\left\{0, \frac{\left(E_p^{i-1} + a(\mathcal{R}_p^i)\right) - \zeta_p^i}{\zeta_p^i}\right\} \tag{9}$$

Minimize Utility Loss: While working towards sustainability and safety goals, we may not be able to ensure maximum utility to customers, however we have to care about the customer utility too at the same time. Thus, instead of a hard constraint of maximum utility (eq-7), we propose to minimize the loss in customer utility while deciding the recommendation at any i^{th} customer login instance.

$$\underset{\mathcal{R}^i}{\arg\min} \left\{1 - NDCG@k(\mathcal{R}^i, i)\right\} \tag{10}$$

Further we reduce the loss minimization objective to a *min-sum* objective in proposition 4.1. This reduction will be helpful while combining this objective with others in eq-8 and eq-9 in order to form a joint optimization problem.

Proposition 4.1. *The objective in equation 10 can be reduced to a min-sum objective.*
$\underset{\mathcal{R}^i}{\arg\min} \left\{ \sum_{p\in P} \frac{V^i(\mathcal{R}^{i*}[1]) - V_p^i}{V^i(\mathcal{R}^{i*}[1])} \times a(\mathcal{R}_p^i)\right\}$
(proof available in extended version of paper at https://bit.ly/3kzBzYz)

4.2 Joint Optimization of Relaxed Objectives

We combine the relaxed objectives: eq-8 for sustainability, eq-9 for safety, and reduced form of eq-10 in proposition-4.1 for customer utility with weights λ_1, λ_2, and $(1-\lambda_1-\lambda_2)$ respectively. We write the joint optimization problem below.

$$\begin{aligned}
\underset{X^i}{\arg\min} \Bigg(& \lambda_1 \sum_{p\in P}\sum_{j=1}^{n} \max\left\{0, \frac{\beta i - \left(E_p^{i-1} + a(j)\right)}{\beta i}\right\} \cdot X_{p,j}^i \\
& + \lambda_2 \sum_{p\in P}\sum_{j=1}^{n} \max\left\{0, \frac{\left(E_p^{i-1} + a(j)\right) - \zeta_p^i}{\zeta_p^i}\right\} \cdot X_{p,j}^i \\
& + (1-\lambda_1-\lambda_2) \sum_{p\in P}\sum_{j=1}^{n} \frac{V^i(\mathcal{R}^{i*}[1]) - V_p^i}{V^i(\mathcal{R}^{i*}[1])} \cdot a(j) \cdot X_{p,j}^i \Bigg) \\
\text{s.t.} & \Big(\sum_{j=1}^{n} X_{p,j}^i = 1, \forall p \in P\Big) \& \Big(\sum_{p\in P} X_{p,j}^i = 1, \forall j \in \{1, 2, \cdots, n\}\Big) \\
& \& \Big(X_{p,j}^i \in \{0,1\}, \forall j \in \{1, 2, \cdots, n\}, p \in P\Big)
\end{aligned} \tag{11}$$

where $X_{p,j}^i$ is a binary indicator variable, $X_{p,j}^i = 1$ represents that business p is assigned j^{th} rank in the recommendation at i^{th} customer login, $\forall j \in \{1, 2, \cdots, n\}, p \in P$. Above optimization problem is an integer linear program which is a discrete optimization problem and computationally heavy to solve. Thus, we plan to reorganize this problem and map it to a matching problem with polynomial time solutions.

4.3 Mapping to Bipartite Min-Cost Perfect Matching

Here, we outline basic details of general bipartite min-cost matching problem and then describe our mapping.

Bipartite Min-Cost Perfect Matching Problem: In a general bipartite minimum-cost perfect matching problem, we are given a complete bipartite graph with two sets of nodes Y, Z s.t. $|Y| = |Z| = n$, and the costs of all the edges between Y, Z: i.e., $c(y, z)$ = cost of edge between $y, z, \forall y \in Y, z \in Z$; all the costs are non-negative: $c(y, z) \geq 0, \forall y, z$. Here, the goal is to find a perfect matching with the minimum cost. A matching is a set of pairwise non-adjacent edges, i.e., no two edges share a common node. A perfect matching is a matching which covers all the nodes in the graph. There can be multiple perfect matchings in this problem, however we need to find the one which costs the least.

We can use a matrix (X) to represent a matching where element $X_{y,z} = 1$ if edge (y, z) is a part of the matching, and $X_{y,z} = 0$ otherwise. Now the min-cost perfect matching problem can be expressed as a discrete optimization problem using X as given below.

$$\begin{aligned}
\underset{X}{\arg\min} & \sum_{y\in Y}\sum_{z\in Z} c(y,z)\cdot X_{y,z} \\
\text{s.t.} & \Big(\sum_{y\in Y} X_{y,z} = 1, \forall z \in Z\Big) \& \Big(\sum_{z\in Z} X_{y,z} = 1, \forall y \in Y\Big) \\
& \& \Big(X_{y,z} \in \{0,1\}, \forall y \in Y, z \in Z\Big)
\end{aligned} \tag{12}$$

The first two constraints ensure that each node from Y and Z is covered by exactly one edge in the solution such that there are no pairwise adjacent edges.

LP Relaxation: Relaxing the third constraint $X_{y,z} \in \{0,1\}$ to $X_{y,z} \geq 0$, converts the problem into a *linear program* (LP) as given below.

$$\begin{aligned}
\underset{X}{\arg\min} & \sum_{y\in Y}\sum_{z\in Z} c(y,z)\cdot X_{y,z} \\
\text{s.t.} & \Big(\sum_{y\in Y} X_{y,z} = 1, \forall z \in Z\Big) \& \Big(\sum_{z\in Z} X_{y,z} = 1, \forall y \in Y\Big) \\
& \& \Big(X_{y,z} \geq 0, \forall y \in Y, z \in Z\Big)
\end{aligned} \tag{13}$$

It turns out that this LP in equation 13 has integer solutions as extreme points (proven in [7, 16, 38] using *Birkhoff-von Neumann theorem*). Thus, we no longer need the integrality constraint (third constraint) of the discrete optimization problem (in equation 12), and instead solve the LP-relaxation of the same (in equation 13) to get to the same solution. Next we discuss how we use this to our advantage in order to efficiently solve our problem.

The Mapping: First, we reorganize our joint optimization problem (eq-11) a bit and write it as below.

$$\begin{aligned}
\underset{X^i}{\arg\min} \sum_{p\in P}\sum_{j=1}^{n} \Bigg(& \lambda_1 \cdot \max\left\{0, \frac{\beta i - \left(E_p^{i-1} + a(j)\right)}{\beta i}\right\} \\
& + \lambda_2 \cdot \max\left\{0, \frac{\left(E_p^{i-1} + a(j)\right) - \zeta_p^i}{\zeta_p^i}\right\} \\
& + (1-\lambda_1-\lambda_2) \cdot \frac{V^i(\mathcal{R}^{i*}[1]) - V_p^i}{V^i(\mathcal{R}^{i*}[1])} \times a(j) \Bigg) \cdot X_{p,j}^i \\
\text{s.t.} & \Big(\sum_{j=1}^{n} X_{p,j}^i = 1, \forall p \in P\Big) \& \Big(\sum_{p\in P} X_{p,j}^i = 1, \forall j \in \{1, 2, \cdots, n\}\Big) \\
& \& \Big(X_{p,j}^i \in \{0,1\}, \forall j \in \{1, 2, \cdots, n\}, p \in P\Big)
\end{aligned} \tag{14}$$

We can see that this is in the same format as the one for matching problem in eq-12. Thus, we can map our problem of finding ranked recommendation $\mathcal{R}^i$ at i^{th} customer login instance to a bipartite min-cost matching problem. In one set of nodes, we have all the businesses (P) and in the other one, we have all the positions/ranks ($[n] = \{1, \cdots, n\}$); there is a cost attached to each edge between P and $[n]$. $c^i(p, j) = \lambda_1 \cdot \max\left\{0, \frac{\beta i - \left(E_p^{i-1} + a(j)\right)}{\beta i}\right\} + \lambda_2 \cdot \max\left\{0, \frac{\left(E_p^{i-1} + a(j)\right) - \zeta_p^i}{\zeta_p^i}\right\} + (1 - \lambda_1 - \lambda_2) \cdot \frac{V^i(\mathcal{R}^{i*}[1]) - V_p^i}{V^i(\mathcal{R}^{i*}[1])} \times a(j)$. The terms multiplied with λ_1, λ_2, $(1 - \lambda_1 - \lambda_2)$ represent the costs which the edge between p, j (i.e., p being assigned j^{th} rank) contributes to exposure deficit objective in eq-8, exposure surplus objective in eq-9, and customer utility objective in proposition-4.1 respectively.

Now, the goal is to find a perfect matching of minimum cost. We can use the LP relaxation (just as done for matching problem in eq-13) to solve this problem (eq-14) in polynomial time. Formally, we write the problem as below.

$$\begin{aligned} &\arg\min_{X^i} \sum_{p \in P} \sum_{j=1}^{n} c^i(p, j) \cdot X_{p,j}^i \\ &\quad \text{s.t.} \Big(\sum_{j=1}^{n} X_{p,j}^i = 1, \forall p \in P \Big) \& \Big(\sum_{p \in P} X_{p,j}^i = 1, \forall j \in \{1, 2, \cdots, n\} \Big) \\ &\qquad \& \Big(X_{p,j}^i \geq 0, \forall j \in \{1, 2, \cdots, n\}, p \in P \Big) \end{aligned} \tag{15}$$

4.4 Approximate Solution with Prefiltering

As the above defined LP operates on the whole set of businesses, huge number of businesses in realtime can be a bottleneck resulting in long processing times. Thus we propose to prefilter the set in the following two ways: (i) top-k^2 businesses based on V^i, which can help in achieving better customer utility, and (ii) k^2 least exposed businesses, which can help in achieving better sustainability and safety. We then merge these two filtered lists to get a smaller set of businesses and run the LP on it.

5 EXPERIMENTAL EVALUATION

Here, we explain the setup and baselines for comparison. Then we detail evaluation metrics and experimental results.

Setting k: We fix $k = 10$ in all the experiments.

Setting β: As there are n businesses, we limit setting β to $\beta_{\max} = \frac{1}{n}$. We vary β from $(0.1 \times \beta_{\max})$ to $(1.0 \times \beta_{\max})$.

Setting λ_1 and λ_2: We vary both λ_1 and λ_2 in the range $[0.1, 0.5]$.

Baselines: We use the following schemes for ranked recommendations at each customer instance (i) as baselines while we refer to our proposed recommendation mechanism as **LP** (as detailed in §4).

(1) Top-k: Recommendation of most relevant k ranked in descending order of relevance V_p^i.

(2) Top-k (Safe): Top-k relevant businesses out of all those which satisfy the safety criterion (eq-6).

(3) Mixed-k: Here, we build a k-sized ranking with top $\left\lfloor \frac{k}{2} \right\rfloor$ in descending order of V_p^i, and the next $(k - \left\lfloor \frac{k}{2} \right\rfloor)$ in ascending order of exposure E_p^{i-1} while ensuring no business is repeated in recommendation.

(4) Mixed-k (Safe): Mixed-k ranking out of all those businesses which satisfy the safety criterion (eq-6).

(5) Poorest-k: Here, we build a k-sized ranking in ascending order of exposure E_p^{i-1} (i.e., k least exposed businesses).

5.1 Evaluation Metrics

We use the following metrics to capture the performance from sustainability, safety and utility standpoints.

A. Metrics for Business Sustainability:

Inequality in Business Exposures (INQ)***:*** We use Gini coefficient [15] to measure inequality in business exposures.

$$INQ = \frac{\sum_{p_1 \in P} \sum_{p_2 \in P} \left| E_{p_1} - E_{p_2} \right|}{2n \sum_{p \in P} E_p} \tag{16}$$

$INQ \in [0, 1]$, and lower value of INQ represents less inequality in business exposures and more business sustainability.

Mean Exposure Loss on Businesses ($ELoss$)***:*** While we guarantee a minimum exposure for all the businesses, there will be some popular businesses who will lose a share of their exposures which they would have received in conventional top-k recommendations. Thus here, we measure the mean exposure loss of all the businesses as defined below.

$$ELoss = \frac{1}{n} \sum_{p \in P} \max\left\{0, \frac{E_p^{\text{top-k}} - E_p}{E_p^{\text{top-k}}}\right\} \tag{17}$$

where $E_p^{\text{top-k}}$ is the exposure received by p in top-k recommendations. $ELoss \in [0, 1]$, and lower $ELoss$ represents lower exposure loss for businesses in comparison to top-k recommendations.

B. Metrics for Customer Safety:

Mean Risk for Customers ($MRisk$)***:*** For safety, we measure the mean chances of customers ending up at an already overexposed business i.e., the mean customer attention directed towards overexposed businesses.

$$MRisk = \frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \sum_{p \in \mathcal{R}^i} a(\mathcal{R}_p^i) \cdot \mathbb{1}_{E_p^{i-1} > \zeta_p^{i-1}} \tag{18}$$

where $\mathbb{1}_{E_p^{i-1} > \zeta_p^{i-1}}$ is 1 if $E_p^{i-1} > \zeta_p^{i-1}$, and 0 otherwise. As $\mathcal{R}_p^i$ is the rank of p in $\mathcal{R}^i$, thus $(a(\mathcal{R}_p^i) \cdot \mathbb{1}_{E_p^{i-1} > \zeta_p^{i-1}})$ becomes the customer attention to p if p is already overexposed. $MRisk \in [0, 1]$. Lower $MRisk$ means better customer safety.

Mean Exposure Surplus ($ESrp$)***:*** Along with the previous metric, we also measure the mean exposure surplus which represents the expected fraction of overexposure of a business.

$$ESrp = \frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \frac{1}{n} \sum_{p \in P} \max\left\{0, \frac{E_p^i - \zeta_p^i}{\zeta_p^i}\right\} \tag{19}$$

$ESrp \in [0, 1]$, and lower $ESrp$ represents less overcrowding and more customer safety.

C. Recommendation Utility: We also look at the mean (μ_ϕ) of customer utilities or recommendation utilities (ϕ defined in section 2.3) over the customer instances. While the top-k recommendations ensure maximum utility ($\phi = 1$), other recommendation mechanisms are desired to have small losses in compasion to that.

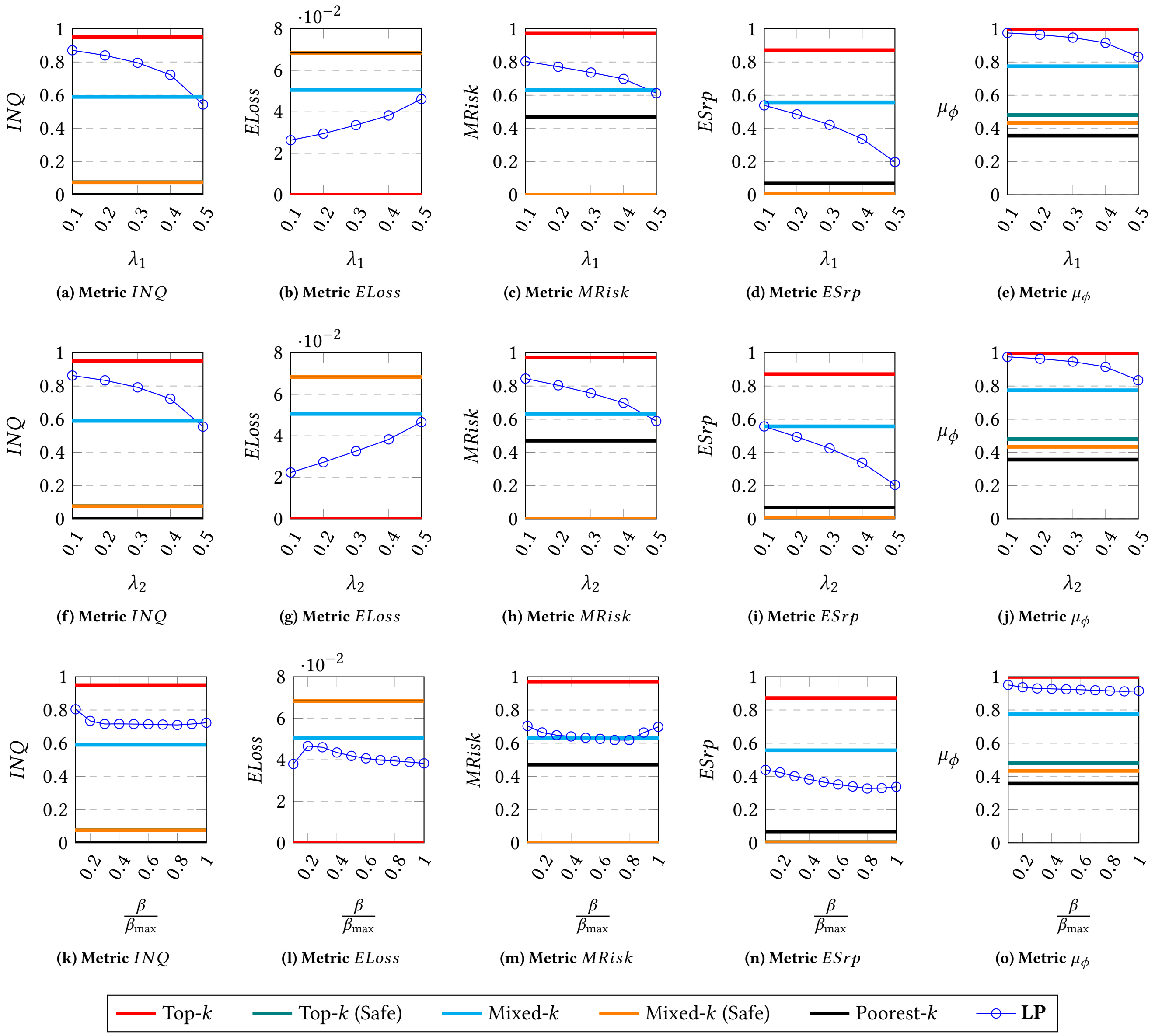


Figure 2: Performances on GL-NYC data. For plots in first row, $\beta = \beta_{\max}$ and $\lambda_2 = 0.4$ (fixed). For plots in second row, $\beta = \beta_{\max}$ and $\lambda_1 = 0.4$ (fixed). For plots in third row, $\lambda_1 = \lambda_2 = 0.4$.

5.2 Experimental Results

We simulate the recommender system as detailed in §3, and run all the baseline mechanisms (listed in §5) along with our proposed LP. Next we calculate the metrics (listed in section 5.1) for each of the methods, and plot them. Even though we observe similar performance patterns in all the datasets, due to space constraints, we show plots only for GL-NYC data in figure-2 in this paper; Results on other datasets can be found in the extended version of the paper available at https://bit.ly/3kzBzYz. We test our LP with different hyperparameter settings in separate simulations. In the first row of figure-2 (2a, 2b, 2c, 2d, 2e), we fix $\beta = \beta_{\max}$, $\lambda_2 = 0.4$, and vary λ_1. In the second row of figure-2 (2f, 2g, 2h, 2i, 2j), we fix $\beta = \beta_{\max}$, $\lambda_1 = 0.4$, and vary λ_2. In the third row of figure-2 (2k, 2l, 2m, 2n, 2o), we fix $\lambda_1 = \lambda_2 = 0.4$, and vary β from $(0.1 \times \beta_{\max})$ to $(1.0 \times \beta_{\max})$. On the other hand, the baselines do not have hyperparameters $(\lambda_1, \lambda_2, \beta)$, thus we plot baseline performances as horizontal straight lines in figure-2 (thus same baseline results in all the three rows).

Performance of Baselines: Even though the conventional top-k recommendation ensures highest customer utility (μ_ϕ in fig-2e), it is very unsuitable from business sustainability and customer safety standpoints, as it causes the highest exposure inequality for businesses (*INQ* in fig-2a), the highest risk for customers (*MRisk* in fig-2c) and the highest amount of exposure surplus (*ESrp* in fig-2d). On the other hand the poorest-k recommendation performs the best in business sustainability with the lowest exposure inequality (*INQ* in fig-2a), and better in customer safety with a lower risk for customers (*MRisk* in fig-2c) and a much lower overexposure (*ESrp* in fig-2d) than top-k; however, the poorest-k performs the worst from customer utility standpoint (μ_ϕ in fig-2e), as it does not take relevance scores into account while deciding the recommendations. In comparison to top-k and poorest-k, the mixed-k recommendation performs fairly good from all the standpoints—reduced exposure inequality than top-k (*INQ* in fig-2a) with smaller exposure loss than poorest-k (*ELoss* in fig-2b), lower customer risk (*MRisk* in fig-2c) and lower overexposure (*ESrp* in fig-2d) than top-k with a much smaller loss in customer utility (μ_ϕ in fig-2e) than poorest-k; this is because mixed-k combines highly relevant businesses along with less exposed businesses in the recommendations thus covering for both customer utility and exposure inequality. The top-k (safe) and mixed-k (safe) recommendations with safety as a hard constraint, perform the best in customer safety with the lowest risk to customers (*MRisk* in fig-2c) and the lowest overexposure (*ESrp* in fig-2d); they also perform very good in business sustainability with very low exposure inequality (*INQ* in fig-2a); however they perform very poorly from utility standpoint with huge loss in customer utility (μ_ϕ in fig-2e) close to that of poorest-k. Out of all the baselines, only mixed-k performs fairly good in all the metrics of interest.

Performance of LP: Observed patterns from the first row of plots in figure-2 suggest that larger settings of λ_1 (i.e., larger weight for sustainability) leads to better business sustainability (lower *INQ* in fig-2a) with marginal loss to previously popular businesses (small *ELoss* in fig-2b), and better customer safety (lower *MRisk*, *ESrp* in fig-2c, 2d) with marginal loss in customer utility (μ_ϕ in fig-2e: less than 20% loss in comparison to top-k). Similarly from the second row of plots in figure-2, we see that increasing λ_2 yields better business sustainability (lower *INQ* in fig-2f) with marginal loss to previously popular businesses (small *ELoss* in fig-2g), and better customer safety (lower *MRisk*, *ESrp* in fig-2h, 2i) with marginal loss in customer utility (μ_ϕ in fig-2j: less than 20% loss in comparison to top-k). From the above observations, we can say that increasing weights for either of business sustainability objective (λ_1) or customer safety objective (λ_2) leads to improvements in both sustainability and safety metrics; the reasons behind this can be: (i) ensuring sustainability through minimum exposure guarantee for every business preferably allots exposure to less exposed businesses ultimately leading to less overexposure and better customer safety, (ii) ensuring safety through capacity-based upper limits on business exposures leads to redistribution of extra exposure of overexposed businesses to less exposed ones, thereby resulting in better sustainability for more number of businesses. In summary, we can say that **business sustainability and customer safety complement each other**. In the third row of figure-2, by increasing β from $(0.1\times\beta_{\max})$ to $(1.0\times\beta_{\max})$, we see better sustainability (fig-2k: decrease in *INQ*), marginal decrease in customer utility (fig-2o: decrease in μ_ϕ), and slightly better safety at first (till $\beta = 0.8\cdot\beta_{\max}$ in fig-2m, 2n); however after $\beta = 0.8\beta_{\max}$, there is a small increase in customer risk and overexposure (*MRisk*, *ESrp* in fig-2m, 2n); this is happening as the minimum exposure guarantee (β) is increased more and more, at some point it grows beyond the upper exposure limit set based on the capacity of some businesses—especially for businesses with low capacity—making the sustainability objective in conflict with safety objective. Thus, the value of β should be set carefully so that it does not come in conflict with the safety objective.

In comparison to the best performing baseline (mixed-k), LP with (λ_1, λ_2) set around $(0.4, 0.4)$, performs better with less overexposure (*ESrp* in fig-2d, 2i) and better customer utility (μ_ϕ in fig-2e, 2j), while it shows similar performances in other metrics. Moreover our proposed LP gives a three-way control over sustainability, safety, and utility objectives which the baselines do not possess; this kind of control can be very useful in the post-pandemic world as the hyperparameters can be set higher or lower according to the realtime peaks or drops in viral infection rates.

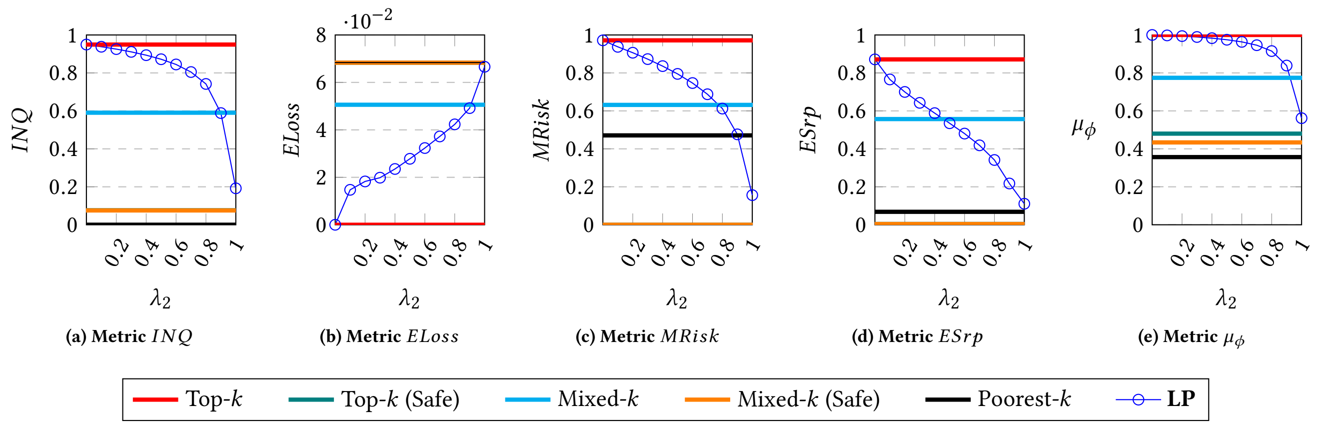
5.3 Turning Crisis into Opportunity

While sustainability is a business-side requirement, both safety and utility are customer-side requirements. Besides, the safety requirement has recently become important because of the pandemic while the sustainability of businesses has been a growing concern for a longer time [8, 29, 36]. Based on one of the important findings in the last section ("*business sustainability and customer safety complement each other*"), we hypothesize that solving only the crisis-inspired safety problem could ultimately solve the long standing business sustainability concern. Thus, we also test our LP with only customer-side objectives (safety and utility) but no sustainability objective (i.e., $\lambda_1 = 0$), and plot the results in figure-3. The results show that increasing the weights of safety objective (λ_2) results in better customer safety (fig-3c, 3d: decrease in *MRisk*, *ESrp*) and also better business sustainability (fig-3a: decrease in *INQ*). In fact, the performance of LP with $(\lambda_1, \lambda_2) = (0, 0.8)$ here, is very similar to that of the best performing LP setting $(\lambda_1, \lambda_2) = (0.4, 0.4)$ in the last section (section 5.2). Thus, in summary we can say that the crisis-inspired safety problem could very well be turned into an opportunity to solve sustainability problem without explicitly addressing it.

6 RELATED WORK

Use of Location in Recommender Systems: Extensive use of ubiquitous location-based services like geo-tagging on social media posts, realtime reviews of establishments on review forums, maps with realtime navigation, etc. have resulted in the generation of huge amount of data on the different user preferences based on location. This has led to a range of new research problems on how to utilise this location-based data to design personalised recommender systems. Recommender systems that consider the location of its users, can have a myriad of applications ranging from recommending nearby restaurants/shops [24, 44] and social events [33], to endorsing neighborhoods to reside [10, 46] or suggesting friends who are nearby [45]. They can be further adapted for trajectory recommendations [47], or recommendation of location sequence for itinerary planning with time/cost optimization [5, 6].

Advancements in Local Recommendations: A key feature which distinguishes location-based recommendation from regular item



(a) Metric INQ (b) Metric $ELoss$ (c) Metric $MRisk$ (d) Metric $ESrp$ (e) Metric μ_ϕ

Figure 3: Performances on GL-NYC with only customer-side objectives (safety and utility), and no sustainability objective i.e., $\lambda_1 = 0$.

recommendation is the location data of source (customers) and destination (establishments/businesses). The approaches used for local recommendation can be considered to be inspired by those in item recommendation literature. Initial work on location-based recommendation have introduced methods to use location data as a personalised filtering criteria in recommendation, using content based filtering [28, 34]. Researchers have also proposed techniques to embed geospatial data into collaborative filtering based recommender systems to suggest nearby places [24, 44]. On the other hand, studies have also proposed hybrid approaches [33, 43] for local recommendations. The availability of geo-tagging features in social media posts on a variety of platforms, has attracted new research initiatives that exploit the social network structure for better design of local recommendations. For example, leveraging social computing techniques to get local experts and then using it for local recommendations [3], social link analysis using graph-based modeling for better estimatation of relevance scores [35, 40], formation of location-based social networks—with individuals connected by the interdependency derived from their locations in the physical world as well as their location-tagged media content—as a solution to local recommendation [4, 9, 26], etc. Moreover, advanced machine learning techniques like deep representation learning for local recommendation [19, 32] and information retrieval techniques like learning-to-rank for ranked recommendation [23] have also been explored for this purpose. However, none of these studies have considered the notions of customer safety and business sustainability which are of primal importance in the post-pandemic world. Besides, these works can easily replace our relevance scoring method (in section 2.4), and serve as inputs to our LP based solution, as all of them ultimately estimate some form of relevance scores.

Sustainability in Recommender Systems: From an orthogonally different research area of algorithmic fairness, few recent works have focussed on the sustainability of businesses in recommender systems [8, 29], recommendation updates [30], etc. in a general online market setup. However, in this paper, we define the notion of business sustainability in local recommendation setup and address it along with the notion of customer safety and utility.

7 CONCLUSION

In this work we formally define timely notions of business sustainability and customer safety in local recommendation, and address them using a novel mapping to min-cost matching problem. Our proposed mechanism is not only computationally efficient, but also easily adaptable as it is independent of the selection of relevance scoring method or any other domain-specific business logic already in place on the platforms. We demonstrate the efficacies of our mechanism through extensive evaluations on gathered datasets. The three-way control over business sustainability, customer safety, and utility goals, can be very useful in the post-pandemic world as the hyperparameters can be set higher or lower according to the need of the situations. Apart from that the idea of capacity-based safety notion, can also be generalized to design safety-aware local recommendations for indoor and outdoor monuments, public parks, etc. As it is said a crisis opens up new avenues, the change in recommendation scheme as well as people's habit would provide more opportunities to local businesses to flourish which perhaps earlier could not survive due to existence of 'popular' outlets.

Acknowledgements: G. K Patro is supported by TCS Research Fellowship. This research was supported in part by a European Research Council (ERC) Advanced Grant for the project "Foundations for Fair Social Computing", funded under the European Unions Horizon 2020 Framework Programme (grant agreement no. 789373).

Reproducibility: The dataset and code base can be found at https://bit.ly/3kzBzYz.

REFERENCES

[1] Aman Agarwal, Ivan Zaitsev, Xuanhui Wang, Cheng Li, Marc Najork, and Thorsten Joachims. 2019. Estimating position bias without intrusive interventions. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*. ACM, 474–482.

[2] David Arthur and Sergei Vassilvitskii. 2006. *k-means++: The advantages of careful seeding*. Technical Report. Stanford.

[3] Jie Bao, Yu Zheng, and Mohamed F Mokbel. 2012. Location-based and preference-aware recommendation using sparse geo-social networking data. In *Proceedings of the 20th international conference on advances in geographic information systems*. 199–208.

[4] Jie Bao, Yu Zheng, David Wilkie, and Mohamed Mokbel. 2015. Recommendations in location-based social networks: a survey. *GeoInformatica* 19, 3 (2015), 525–565.

[5] Ramesh Baral, SS Iyengar, Tao Li, and N Balakrishnan. 2018. Close: contextualized location sequence recommender. In *Proceedings of the 12th ACM conference on recommender systems*. 470–474.

[6] Ramesh Baral, SS Iyengar, Tao Li, and XiaoLong Zhu. 2018. HiCaPS: Hierarchical contextual poi sequence recommender. In *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. 436–439.

[7] Garrett Birkhoff. 1946. Three observations on linear algebra. *Univ. Nac. Tucumán. Revista A.* 5 (1946), 147–151.

[8] Robin Burke. 2017. Multisided fairness for recommendation. *arXiv preprint arXiv:1707.00093* (2017).

[9] Chen Cheng, Haiqin Yang, Michael R Lyu, and Irwin King. 2013. Where you like to go next: Successive point-of-interest recommendation. In *Twenty-Third international joint conference on Artificial Intelligence*.

[10] Justin Cranshaw, Raz Schwartz, Jason Hong, and Norman Sadeh. 2012. The livehoods project: Utilizing social media to understand the dynamics of a city. In *Sixth International AAAI Conference on Weblogs and Social Media*.

[11] Nick Craswell, Onno Zoeter, Michael Taylor, and Bill Ramsey. 2008. An experimental comparison of click position-bias models. In *Proceedings of the 2008 international conference on web search and data mining*. ACM, 87–94.

[12] Christian Desrosiers and George Karypis. 2011. A comprehensive survey of neighborhood-based recommendation methods. In *Recommender systems handbook*. Springer, 107–144.

[13] Niklas Engbom and Christian Moser. 2018. Earnings inequality and the minimum wage: Evidence from Brazil. *Federal Reserve Bank of Minneapolis-Opportunity and Inclusive Growth Institute Working Paper* 7 (2018), 18–50.

[14] Armin Falk, Ernst Fehr, and Christian Zehnder. 2006. Fairness perceptions and reservation wages—the behavioral effects of minimum wage laws. *The Quarterly Journal of Economics* 121, 4 (2006), 1347–1381.

[15] Corrado Gini. 1921. Measurement of inequality of incomes. *The Economic Journal* 31, 121 (1921), 124–126.

[16] Michel X Goemans and David P Williamson. 1997. The primal-dual method for approximation algorithms and its application to network design problems. *Approximation algorithms for NP-hard problems* (1997), 144–191.

[17] Google. 2020. Google Places API. https://developers.google.com/places/web-service/intro.

[18] David A Green and Kathryn Harrison. 2010. *Minimum wage setting and standards of fairness*. Technical Report. IFS Working Papers.

[19] Zhenhua Huang, Xiaolong Lin, Hai Liu, Bo Zhang, Yunwen Chen, and Yong Tang. 2020. Deep Representation Learning for Location-Based Recommendation. *IEEE Transactions on Computational Social Systems* (2020).

[20] Kalervo Järvelin and Jaana Kekäläinen. 2002. Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems (TOIS)* 20, 4 (2002), 422–446.

[21] Yehuda Koren, Robert Bell, and Chris Volinsky. [n.d.]. Matrix factorization techniques for recommender systems. *Computer 42.8 (2009): 30-37* ([n. d.]).

[22] Hang Li. 2011. Learning to rank for information retrieval and natural language processing. *Synthesis Lectures on Human Language Technologies* 4, 1 (2011), 1–113.

[23] Jian Li, Guanjun Liu, Chungang Yan, and Changjun Jiang. 2019. LORI: a learning-to-rank-based integration method of location recommendation. *IEEE Transactions on Computational Social Systems* 6, 3 (2019), 430–440.

[24] Defu Lian, Cong Zhao, Xing Xie, Guangzhong Sun, Enhong Chen, and Yong Rui. 2014. GeoMF: joint geographical modeling and matrix factorization for point-of-interest recommendation. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. 831–840.

[25] Carl Lin, Myeong-Su Yun, et al. 2016. The effects of the minimum wage on earnings inequality: Evidence from China. *Income Inequality Around the World* 44 (2016).

[26] Bin Liu, Yanjie Fu, Zijun Yao, and Hui Xiong. 2013. Learning geographical preferences for point-of-interest recommendation. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1043–1051.

[27] NDTV. 2020. Coronavirus Apps: Every App the Central Government And States Have Deployed to Track COVID-19. https://gadgets.ndtv.com/apps/features/central-state-governments-launch-coronavirus-mobile-app-list-2204286. (2020).

[28] Moon-Hee Park, Jin-Hyuk Hong, and Sung-Bae Cho. 2007. Location-based recommendation system using bayesian user's preference model in mobile devices. In *International conference on ubiquitous intelligence and computing*. Springer, 1130–1139.

[29] Gourab K Patro, Arpita Biswas, Niloy Ganguly, Krishna P Gummadi, and Abhijnan Chakraborty. 2020. FairRec: Two-Sided Fairness for Personalized Recommendations in Two-Sided Platforms. In *Proceedings of The Web Conference 2020*. 1194–1204.

[30] Gourab K Patro, Abhijnan Chakraborty, Niloy Ganguly, and Krishna P Gummadi. 2020. Incremental Fairness in Two-Sided Market Platforms: On Smoothly Updating Recommendations. *AAAI, Feb* (2020).

[31] Robert Pollin, Mark Brenner, Stephanie Luce, and Jeannette Wicks-Lim. 2008. *A measure of fairness: The economics of living wages and minimum wages in the United States*. Cornell University Press.

[32] Tieyun Qian, Bei Liu, Quoc Viet Hung Nguyen, and Hongzhi Yin. 2019. Spatiotemporal representation learning for translation-based POI recommendation. *ACM Transactions on Information Systems (TOIS)* 37, 2 (2019), 1–24.

[33] Daniele Quercia, Neal Lathia, Francesco Calabrese, Giusy Di Lorenzo, and Jon Crowcroft. 2010. Recommending social events from mobile phone location data. In *2010 IEEE international conference on data mining*. IEEE, 971–976.

[34] Lakshmish Ramaswamy, P Deepak, Ramana Polavarapu, Kutila Gunasekera, Dinesh Garg, Karthik Visweswariah, and Shivkumar Kalyanaraman. 2009. Caesar: A context-aware, social recommender system for low-end mobile devices. In *2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware*. IEEE, 338–347.

[35] Rudy Raymond, Takamitsu Sugiura, and Kota Tsubouchi. 2011. Location recommendation based on location history and spatio-temporal correlations for an on-demand bus system. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. 377–380.

[36] Slate.com. 2019. Billion Dollar Bully Highlights Why Yelp Feels Unfair. https://slate.com/technology/2019/06/billion-dollar-bully-documentary-yelp.html. (2019).

[37] The Verge. 2020. Apple and Google are building a coronavirus tracking system into iOS and Android. https://www.theverge.com/2020/4/10/21216484/google-apple-coronavirus-contract-tracing-bluetooth-location-tracking-data-app. (2020).

[38] John Von Neumann. 1953. A certain zero-sum two-person game equivalent to the optimal assignment problem. *Contributions to the Theory of Games* 2, 0 (1953), 5–12.

[39] Think with Google. 2016. How Mobile Search Connects Consumers to Stores. https://www.thinkwithgoogle.com/consumer-insights/mobile-search-trends-consumers-to-stores/. (2016).

[40] Min Xie, Hongzhi Yin, Hao Wang, Fanjiang Xu, Weitong Chen, and Sen Wang. 2016. Learning graph-based poi embedding for location-based recommendation. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. 15–24.

[41] Dingqi Yang, Bingqing Qu, Jie Yang, and Philippe Cudre-Mauroux. 2019. Revisiting user mobility and social relationships in lbsns: A hypergraph embedding approach. In *The World Wide Web Conference*. 2147–2157.

[42] Dingqi Yang, Daqing Zhang, Vincent W Zheng, and Zhiyong Yu. 2014. Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 45, 1 (2014), 129–142.

[43] Mao Ye, Peifeng Yin, and Wang-Chien Lee. 2010. Location recommendation for location-based social networks. In *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems*. 458–461.

[44] Mao Ye, Peifeng Yin, Wang-Chien Lee, and Dik-Lun Lee. 2011. Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. 325–334.

[45] Fei Yu, Nan Che, Zhijun Li, Kai Li, and Shouxu Jiang. 2017. Friend recommendation considering preference coverage in location-based social networks. In *Pacific-Asia conference on knowledge discovery and data mining*. Springer, 91–105.

[46] Amy X Zhang, Anastasios Noulas, Salvatore Scellato, and Cecilia Mascolo. 2013. Hoodsquare: Modeling and recommending neighborhoods in location-based social networks. In *2013 International Conference on Social Computing*. IEEE, 69–74.

[47] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. 2009. Mining interesting locations and travel sequences from GPS trajectories. In *Proceedings of the 18th international conference on World wide web*. 791–800.