Milan's Diverse Mosaic: Urbanization and Foreign Relations Shaping Ethnic Plurality

İrem Betül Koçak*, and Oğuz Yücel*
kocaki@itu.edu.tr, yucelog@itu.edu.tr
* Management Engineering Department, Istanbul Technical University, 34469 Maçka, Istanbul, Turkey.

I. Introduction

Cities are built in areas with surplus resources, attracting people from rural areas to move and settle there. Consequently, these densely populated areas foster a culture of coexisting as strangers with different backgrounds, ethnicity and cultures. In this study, we utilize Call Detail Records data from the Telecom Italia Dataset[1], which includes the total number of calls, SMS, and Internet calls originating from square IDs within the Milano grid map. We examine the country codes present in the data to determine which countries are most frequently communicated with using these various telecommunication methods. Our aim is to examine how the historical roots, style of urbanization, and the relationship with foreigners impact the cultural and ethnic fabric of Milan, the financial capital of Italy. We also seek to understand the extent of ethnical diversity in the city.

II. Methodology

There are different metrics to measure the diversity in an area and entropy index is one of them. The entropy index is commonly referred to as the Shannon index because it is related to information theory [1]. It reaches its maximum, log N, when all subgroups have the same proportion [2] as in the equation:

$$H = -\sum_{i=1}^{N} p_i * log(p_i) \tag{1}$$

In this study, the entropy index, total call, and total SMS amounts were used to calculate and analyze the diversity within cells in detail. Three distinct locations in the city were selected to examine how entropy changes throughout the day and to assess the factors that may lead to diversity in these areas.

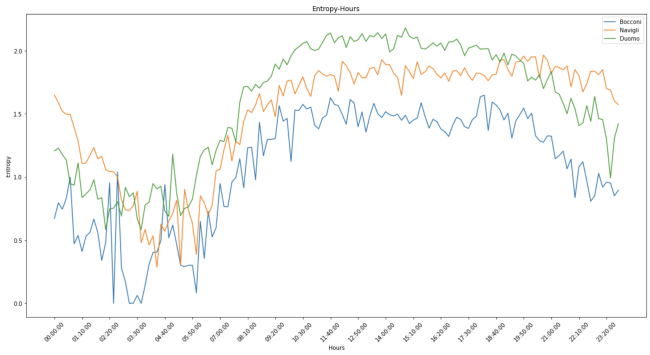

Fig. 1. Entropy changes in three different areas.

As mentioned earlier, we highlighted Milan as one of the prominent cities in Europe in terms of education, tourism, and entertainment. This study examines the entropy values of three distinct areas in Milan: Bocconi University, Navigli, and Duomo Square. Bocconi University shows lower diversity during nighttime but increases in the morning. Duomo Square consistently exhibits high entropy values due to its tourist attraction, while Navigli maintains high entropy values as a vibrant nightlife hub. Duomo Square generally has the highest daily entropy, while Bocconi University has the lowest. This analysis reveals the temporal variations and factors influencing diversity in these areas, providing insights into Milan's education, tourism, and entertainment dynamics.

[1]Details can be found at http://dx.doi.org/10.7910/DVN/UTLAHU

III. Segregation Analysis

We create maps of dominant countries, referring to the most frequently communicated countries, in these bits of Milan grid for two different uses of the city: one for private use during the hours when people are usually sleeping or at home, and the other for public use when people are primarily working or socializing, and tourists are typically exploring the city. We examine the dominance map for these two different uses in Figure 2 and Figure 3 which represents each country with different colors.

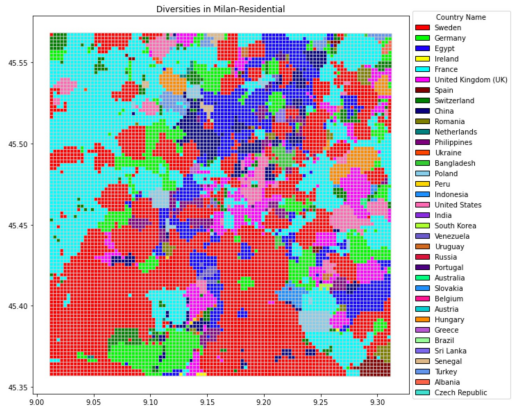

Fig. 2. Milan map with the dominant countries on residential use hours.

We assume that there are four important reasons for communication with a country can dominate a region in the city: People from these countries can emigrate to Italy and communicate with their home countries, people from Milan can migrate to these countries and communicate with their relatives from Milan, tourists can communicate with their home countries, and people communicate internationally for their work. We will consider these four main reasons and leave out the possible other reasons when we talk about dominant countries. Also, we need to consider the communication reason for work less in the residential use hours.

Before examining the dominant countries in particular regions and their relationship with Italy, it is important to discuss the city structure and urbanization of Milan. According to Karaulan [3], the urbanized portion of Milan is located in the middle of the grid, while the peripheral areas are primarily used for industry and

agriculture. Consequently, the peripheral areas cannot be classified as highly urbanized. At first glance, Figure 2 and Figure 3 suggests that the central areas of Milan are more diverse compared to the peripheral area of the city, and this observation supports the acclaimed relationship with urbanization and diversity.

We can also see that there are more countries that could become dominant in their regions in the times of use of residential in certain areas. We can speculate that people tend to group in their housing preferences, so people related to certain countries tend to cluster together and segregate with others more in their residence choice as Schelling claimed [4].

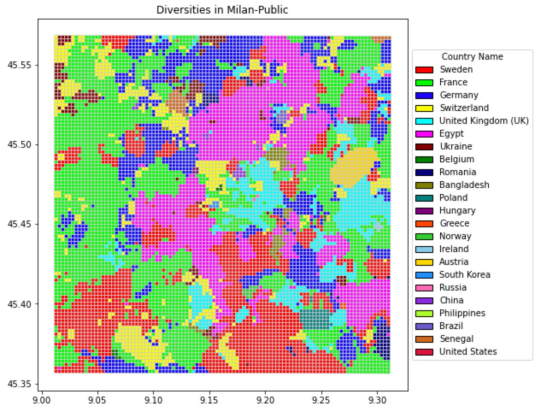

Fig. 3. Milan map with the dominant countries on public use hours.

We observe from both figures that the countries of the European Union (EU) occupy the top five positions. Italy stands out as one of the countries with the highest number of tourist arrivals worldwide, ranking as the third largest among European countries in 2013, according to the UNWTO. Milan, renowned for its historical and tourist attractions, emerges as one of the most alluring cities in Italy. So, there exist numerous factors that attract European visitors to Milan. Notably, one plausible explanation is the significant Italian diaspora in these dominant countries, as reported by the EUROSTAT [5]. Additionally, the facilitated regulations and extensive inter-connectivity between EU nations and Italy further support the presence of these countries in the top five rankings.

In Figures 2 and 3, Egyptian dominance in the urban area is evident. This is a result of Egyptians migrating to European countries like Italy and France in search of better opportunities due to economic challenges and limited job prospects in Egypt. Milan has become a significant destination for Egyptian migrants, who have secured employment in diverse establishments.

Another important aspect is to identify which countries tend to have close coexistence. Both figures demonstrate that regions with similar communication patterns are generally located in close proximity to each other. Areas that engage in communication with EU member states also exhibit spatial proximity. We also see that France often share borders with other countries where French is also spoken. This suggests that proximity practices are established not only through national identity and economic agreements but also through shared language.

In general, these interpretations based on dominance can be further expanded and detailed through second or third majority/minority mapping, aiding our understanding of Milan's ethnic and cultural life and how this situation of urban dominance translates into daily life and urban discrimination.

IV. Attraction Point Analysis

In the second phase of the study, we will examine whether the diversity in Milan contributes to the attractiveness of certain locations within the city. The Tripadvisor dataset from November 2013 was utilized to determine the attractiveness of cells. User opinions were collected to label certain points as attractive. Four different supervised learning methods were employed to predict whether regions are attractive points or not, by using variables from the CDR dataset and entropy. The data was analyzed to generate predictions for the respective regions.

Method	Accuracy	Precision	Recall	F1-Score
Decision Tree (DT)	0.891	0.564	0.373	0.449
Artificial Neural network (ANN)	0.909	0.674	0.492	0.569
Logistic Regression (LR)	0.911	0.852	0.365	0.511
Linear Discriminant Analysis (LDA)	0.911	0.744	0.460	0.569

TABLE I
Classification metrics

The table summarizes the performance metrics of four supervised learning methods: Decision Tree (DT) with 3 maximum-depth, Artificial Neural Network (ANN), Logistic Regression (LR), and Linear Discriminant Analysis (LDA). The accuracy values for all methods are relatively high, ranging from 0.891 to 0.911, indicating overall successful predictions. Precision scores vary across the methods, with LR and ANN demonstrating higher precision values compared to DT and LDA. The recall and F1-scores also exhibit variations among the methods, indicating differences in their abilities to correctly identify positive instances. Overall, ANN and LR show more balanced performance in terms of precision, recall, and F1-score, suggesting their effectiveness in predicting the attractiveness of regions using the given variables and entropy.

V. Conclusion

This study utilizes mobile phone data to gain insights into the nationalities of migrants, tourists, and diaspora in Milan. It highlights the influence of both historical and cultural factors in explaining this phenomenon. Unlike previous studies focusing on diversity density, this research explores the significance of diversity and segregation created by specific countries in Milan. By employing four different supervised learning methods, the study determines the presence of attractive points in different locations. The findings provide valuable patterns and dynamics within Milan's diverse population, offering a broader understanding of diversity in the city.

References

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[2] J. Iceland, "The multigroup entropy index (also known as theil's h or the information theory index)," *US Census Bureau. Retrieved July*, vol. 31, no. 2006, p. 60, 2004.

[3] D. S. Karaulan and A. S. Kubat, "Analyzing fringe belt phenomenon in the historico–geographical structure of milan, italy," *ICONARP International Journal of Architecture and Planning*, 2018.

[4] T. C. Schelling, "Dynamic models of segregation," *Journal of mathematical sociology*, vol. 1, no. 2, pp. 143–186, 1971.

[5] S. Marino, V. Corbanese, and G. Rosas, "Inclusion of migrants in the labour market a comparative analysis of italy and other eu countries."