Mobile phone data statistics as a dynamic proxy indicator in assessing regional economic activity and human commuting patterns

Irina Arhipova1 | Gundars Berzins1 | Edgars Brekis1 | Juris Binde2 | Martins Opmanis3 | Aldis Erglis1 | Evija Ansonska1

1The Faculty of Business, Management and Economics, University of Latvia, Riga, Latvia
2Latvian Mobile Telephone, Riga, Latvia
3President, Chairman of the Management Board, The Institute of Mathematics and Computer Science, University of Latvia, Riga, Latvia

Correspondence
Irina Arhipova, The Faculty of Business, Management and Economics, University of Latvia, 5 Aspazijas Boulevard, Riga, Latvia.
Email: irina.arhipovasalajeva@gmail.com

Funding Information
The University of Latvia and the Ministry of Environmental Protection and Regional Development of the Republic of Latvia, Grant/Award Number: IL/25/2019; The University of Latvia and LMT Ltd., Grant/Award Number: AAP2016/B089

Abstract
Various studies demonstrate that data on mobile phone use are useful when analysing problems in the fields of human activity or population dynamics, including tourism, transportation planning, public administration, etc. However, one of the biggest challenges is related to the restrictions contained in the General Data Protection Regulation that force the use of statistics about mobile operator client activities instead of allowing the analysis of mobile operator data. Therefore, a data analytics approach that does not involve information on the mobility of particular persons was developed, providing economically relevant data on aggregate mobility while protecting personal data. The activity data aggregation was conducted at 15-min intervals in the area of each cellular base station; “activity” is defined as the number of outgoing and incoming calls and sent and received text messages (short message service or SMS) and, in some instances, as the count of unique users. The case study examines all of Latvia’s municipalities, analysing the economic activity level in each municipality in comparison to the mobile phone activity in three periods: 2015–2016, 2017, and 2018. It was concluded that the economic activity in municipalities can be estimated, and positive dynamics of regional development have been detected. Such data and the data analytics method, which provides an understanding of how economic activities evolve in real time in particular locations and economic activity centres, can improve regional development planning and plan implementation. In order to assess which are the centres of economic activity in each municipality and its sphere of influence, the patterns of human commuting and fluctuations of internal activity on workdays and weekends/holidays in 2017–2018 were determined. In general, there is a shortage of reliable data on human commuting within Latvia and its specific regions; therefore, the method described here provides a practical tool for regional governments to keep track of strategy implementation and for strategic gap analysis.

KEYWORDS
commuting, efficiency, principal component analysis
Mobile phones have become an integral part of modern working life and are much more than just tools for information exchange. Unlike landline phones, which can only be used to get the list of recipients and for an analysis of the frequency and duration of calls made, mobile phones can also provide information on the movements of the phone owner over certain periods of time. Investigating particular mobile phone activities (the facts themselves not their content) can provide an insight into the mobility of the population and its economic activity. The conclusions obtained may be useful for decision-making concerning regional development and could serve as metrics characterizing the national economy. Mobile data analysis is an authoritative source of information for problem solving in the fields of human activity recognition, population dynamics, tourism, transport planning, traffic measurement, and public administration. Previously, mobile positioning data have been analysed in the context of residents’ movements (Ahas, Aasa, Silm, & Tiru, 2010; Zonghao, Dongyuan, & Zhengyu, 2013), human home-work commuting (Kung, Greco, Sobolevsky, & Ratti, 2014), automatic recognition of population activities (Chetty, White, & Akther, 2015; Lee & Cho, 2014), estimation of human trajectories (Hoteit, Secci, Sobolevsky, Ratti, & Pujolle, 2014; Larijani, Olteanu-Raimond, Perret, Bredif, & Ziemlicki, 2015; Liu, Janssens, Wets, & Cools, 2013; Zilske & Nagel, 2015) and flows (Balzotti, Bragagnini, Briani, & Cristiani, 2018), as well as patterns of population dynamics (Deville et al., 2014; Trasarti et al., 2015).

The identification of tourists’ destinations (Alexander, Jiang, Murga, & Gonzalez, 2015; Raun, Ahas, & Tiru, 2016), seasonal patterns (Ahas, Aasa, Mark, Pae, & Kull, 2007; Phithakkitnukoon et al., 2015), and travellers’ preferences (Y. Wang et al., 2018) and behaviour (Z. Wang et al., 2018), as well as evaluations of the tourism sector (Ahas, Aasa, Roose, Mark, & Silm, 2008; Kuusik, Nilbe, Mehine, & Ahas, 2014), travel flow (Ni, Wang, & Chen, 2018), tourist movement patterns (Zhao, Lu, Liu, Lin, & An, 2018), and trip modelling (Bwambale, Choudhury, & Hess, 2019), and analysis of the number of travellers (Sørensen et al., 2018) and passenger demands (Hatziioannidu & Polydoropoulou, 2017) are popular investigation topics.

Mobile phone data have been used for transport planning (Elías, Nadler, Stehno, Krosche, & Lindorfer, 2016; Liu et al., 2014), traffic measurement (Dong et al., 2015; Hongyan & Fasheng, 2013; Steenbruggen, Tranos, & Rietveld, 2016) and modelling (Oliveira, Viana, Naveen, & Sarrate, 2017), trajectory evaluation (Bonnel, Hombourger, Olteanu-Raimond, & Smoreda, 2015; Chen, Bian, & Ma, 2014), and travel time predictions (Woodard et al., 2017). The general problems in urban planning (Jonge, Pelt, & Roos, 2012; Ricciato, Widhalm, Pantisano, & Craglia, 2017) and analysis (Lee et al., 2018), land use (Rios & Muñoz, 2017), and smart city development (Steenbruggen, Tranos, & Nijkamp, 2015) can all be addressed with the help of mobile phone data analysis.

As the basic unit of the mobile network infrastructure is a cell with its own cellular base station, at the beginning of each call, a specific cellular base station determines the location of the mobile phone. If, during a call, the mobile phone is moved beyond the limits of a particular cell, switching to another cellular base station takes place. If a person is sufficiently mobile, then a large number of cellular base stations can be used to conduct or initialize calls.

Although more than one cellular base station may be used during a call, enough information can be obtained by analysing only where the calls have started. If there are several cellular base stations used for initializing calls during the day, then the location of stations quite accurately reveals the habits of its owner. For example, if conversations in the mornings and evenings are initiated from one cellular base station, but from another during the working day, it may be assumed that these conversations geographically identify the person’s place of work and residence.

The information collected by the cellular base stations on the calls initiated and/or text messages (short message service or SMS) sent can also be studied. Changes in mobile activities during the day (week, month) can provide relevant information on the habits of the region’s population.

Within the framework of the study, it was intended to process data from the mobile telecommunications operator, Latvian Mobile Telephone (LMT), to identify factors influencing society and the national economy. An analysis of the mobile phone data statistics makes it possible to evaluate important societal processes. In the original model, the plan was to study all mobile calls, anonymizing only the number, allowing the researchers to trace the activities of each particular mobile phone user anonymously.

The General Data Protection Regulation (GDPR), determined by the Law on Personal Data Processing, entered into force on May 25, 2018, and provides uniform rules for the protection of personal data throughout the European Union (EU). The GDPR applies to any company, entity, or organization that processes or stores data from identifiable individuals living in the EU (European Parliament, 2016). Taking into account the requirements of the GDPR within the framework of this study meant that individual call switching and corresponding activity were not accessed. The researchers had no access to information about who exactly was calling and where the caller moved; that is, only information on overall call numbers and how many unique users there were connected to a particular mobile base station was provided. It was enough to identify the cellular base stations that ensure connections to the network.

The data of each person’s activity was aggregated in the area of each cellular base station at 15-min intervals, where the activity (calls and SMS) might be outgoing as well as incoming. Additionally, the number of unique users during each interval was counted. The latter provides an insight into the average activity in the area covered by a particular base station. In this way, the requirements of personal data protection were respected. Undoubtedly, a significant part of the original data was lost, but the available data can be used

1
to explore the mobility of people and their everyday habits (such as leaving home in the mornings and returning home in the evenings); to explore social interactions. Mobile phone usage models vary significantly depending on age, education level, and gender. Knowing the behaviour patterns of mobile phone users, it is possible to estimate the share of different social groups in the region, as well as to observe changes in their behaviour patterns and habits (or the proportions of various groups); and to forecast economic activity. Mobile service cost dynamics can quite accurately predict upcoming crises.

Call detail record (CDR) is a digital data recorder used in telephone communications and in other telecommunications equipment, which involves phone conversations or other telecommunications transactions (e.g., text messages) that are transmitted by the device. CDR contains call time, duration, call status, and caller and calling subscriber numbers. The CDR does not contain information on the content of the conversation or SMS. In the previous research, it was concluded that mobile phone data is suitable and updatable for the Latvian regional business index development (Arhipova, Berzins, Brekis, Kravcova, & Binde, 2017).

The regions with similar economic activity patterns were identified using mobile communication data, and it was concluded that mobile phone activities have a statistically significant relationship to regional economic activity such as the gross domestic product, the number of economically active enterprises per thousand inhabitants, municipal budget expenditures, etc. (Arhipova et al., 2019). As a result, the hypothesis that municipalities and regions with lower call activity have lower economic activity compared with other regions with higher call activity could not be rejected.

In order to determine the economic centres of municipalities and their sphere of influence, the patterns of human commuting and fluctuations of internal activity on working days and holidays in 2017–2018 were determined.

The objective of this paper is to develop a method for the assessment of the economic activity and human commuting patterns in any particular region, based on mobile phone data statistics.

At first, using previous research results, an analysis of the CDR data was made, paying particular attention to the seasonality effect on mobile phone activities in municipalities. Second, the municipalities were grouped into categories based on their inferred economic activity during three periods: 2015–2016, 2017, and 2018. Third, the municipalities’ economic activity efficiency and their change dynamics were estimated. Finally, the region’s human commuting patterns were estimated, and conclusions about the regional economic development tendencies in Latvia were made.

The use of virtual reality analytics is one of the ways to improve the business efficiency through understanding processes and data (Chang, 2018). Analytics and visualization require advanced techniques for processing, analysis, and fusion so that it can be effective (Sun, Chang, Yanga, & Liao, 2018; Lee et al., 2018). The use of mobile phone apps has been extensive, and its services can provide true values for different disciplines (Sun, Xiong, & Chang, 2019).

Our work is relevant to data management and knowledge representation, empirical and observational studies, as well as real-world applications using interactive computational visual analytics.

2 | DATA AND METHODS

The CDR, which is automatically generated by the mobile network operator initializing a mobile phone call or text message, contains information on the cellular base stations providing the connection and calling side. Because the coordinates of all cellular base stations are known, the location of a person at the beginning of the call can be determined with appropriate precision. Each database entry includes the following parameters: the total number of calls and text messages, the total number of unique users, the date and the 15-min interval, and the antenna identifier of the mobile network cellular base station and its coordinates.

2.1 | Data

The LMT database used in the current case study consists of the roaming data of the mobile phone call activities from July 25, 2015 to December 31, 2018, altogether 148,911,360 CDRs from 1,235 cellular base stations for 1,256 days with data grouping in 15-min time intervals (chosen by the LMT data provider, taking into the account their own data security considerations). The distribution of network cellular base stations in Latvia’s municipalities and regions was obtained using their geographical coordinates. The administrative division of Latvia includes 110 municipalities and 9 cities (Riga, Jekabpils, Jelgava, Jurmala, Ventspils, Liepaja, Daugavpils, Rezekne, and Valmiera), but for statistical and planning purposes, the country is divided into six larger statistical regions: Kurzeme, Latgale, Pōrgā, Riga, Vidzeme, and Zemgale. The data analysis shows differences in the intensity of call activity between workdays and weekends/holidays, characterizing the economic activity of the area (Arhipova et al., 2017). At the same time, the peak of the call activity was observed at noon on all days (Arhipova et al., 2019), as well the seasonality effect in summer and winter holidays.
In Figure 1, the total number of calls and text messages for three cities—Jelgava, Jurmala, and Ventspils—is shown from August 2015 to December 2018. The highest call activity during summer time was detected in the seaside town of Jurmala, with Jelgava having the lowest call activity, but in Ventspils, mobile call activity did not have a strong seasonal effect. All cities have a seasonal effect in December due to the winter holidays (Christmas and New Year’s Eve).

To group municipalities and cities by their economic activity, a principal component analysis (PCA) was used to determine similarities between municipalities in terms of the mobile phone call activities and to group them into the respective categories.

2.2 | Grouping municipalities by similar economic activity

Before grouping the municipalities into categories based on their economic activity, 119 variables (as the linear combination of the total number of mobile phone activities and the total number of unique users for all municipalities) were developed, depending on the day of the week during the 2015–2018 time periods. The PCA was applied using a Varimax rotation separately for the three periods: 2015–2016, 2017, and 2018. To check the adequacy quality of the applied PCA, the Kaiser-Meyer-Olkin (KMO) test for sampling adequacy was used. The results show that KMO values for the time period analysed are between 0.8 and 1, which indicates that the sampling is adequate.

In the first time period, 67.6% of the total variance is described by the first two principal components (PC), and the KMO value equals 0.990, where the first PC has high values on workdays for municipalities with higher economic activity, but the second PC has high values during weekends/holidays for municipalities with lower economic activity (Arhipova et al., 2019).

For 2017, the results of the applied PCA show that 71.0% of the total variance is described by the first two principal components with a KMO value of 0.987, but for 2018, it is 77.7% with a KMO value of 0.988.

To find out the interpretation of principal components, their average values were calculated for weekdays (Figure 2) and months (Figure 3).

It can be concluded that the first PC has the highest values on workdays and lower values on weekends/holidays and during summer months. In contrast, the second PC has lower values on workdays and higher values on weekends/holidays and during summer months.
The component loadings, which are the correlations of the 119 observed variables with the first two principal components, were used to interpret the meaning of the components. It is hypothesized that municipalities with higher economic activity correlate highly with the first PC, but those with lower economic activity correlate highly with the second PC.

The distribution of municipalities by economic activity can be shown using the loading plot of an orthogonal solution. Municipalities in Latvia can be divided into eight groups according to their economic activity. This enables an understanding of the profile of each municipality, depending on the economic activity on workdays and weekends/holidays, as well as during different seasons.

The summary of the proposed eight groups of municipalities, using the two component loadings or correlation coefficients, is shown in Table 1.

In order to evaluate the effectiveness of the economic activity strategy chosen by each municipality, 40% to 100% efficiency curves were constructed. It is necessary to stress that the amount of variance in each variable explained by the principal components or the component communalities is computed by taking the sum of the squared loadings for that variable, where the component loadings are the correlations of the observed variables with the principal component.

Therefore, the efficiency criterion $EC$ was calculated according to the Formula (1):

$$EC = \sqrt{r_{1n}^2 + r_{2n}^2},$$

where $r_{1n}$ is the correlation coefficient of the observed $n$th variable (linear combination of the total number of mobile phone activities and the total number of unique users in the $n$th county) with the first principal component, but $r_{2n}$ is the correlation coefficient of the observed $n$th variable with the second principal component.
2.3 Human commuting analysis for regional economic activity evaluation

Human commuting can be an additional factor in an analysis of regional economic activity. Human commuting analysis indicates the direction taken and the approximate volume of mobile activity across a certain territory.

Analytics and visualizations can provide true values and real-world contributions to big data computing that map well with geographic information systems (Chang, 2017). To discover the population’s commuting patterns and directions, having only aggregated CDR data from mobile cellular base stations, the Monge–Kantorovich mass movement model (Balzotti et al., 2018) was used. Wasserstein distance calculations provide approximate indications on human commuting patterns across a certain territory in different time periods, allowing an analysis of different times of the day and days of the week and enabling the discovery of seasonality effects on commuting throughout the year. The most important time periods for analysis are the time of the day and the day of the week. Empirical analysis shows that the most significant outgoing activity takes place in the morning when people go to work or perform other duties and in the afternoon when they return home. This behaviour pattern differs between workdays and weekends, in the summer and in the winter. Therefore, the time and the day must be carefully selected for analysis to discover the real commuting patterns.

To analyse the commuting patterns related to the regional economic self-sufficiency of an administrative territory, two time periods were chosen as a result of the empirical analysis. The interval between 7 a.m. and 9 a.m. captures an indication of outgoing commuting (going to work), and the interval between 5 p.m. and 7 p.m. captures an indication of incoming commuting (returning home). According to the data analysis, these 3-hr intervals represent most of the activity related to outgoing and incoming commuting. The careful selection of time intervals is crucial for the CDR weekday activity analysis.

To perform a Wasserstein distance calculation, a territory must be divided into regular quadrants (not necessarily corresponding with the administrative territory boundaries). In Latvia’s case, $5 \times 5$ km$^2$ or $3.1 \times 3.1$ sq mi were used in the calculations. With the use of the GPS coordinates, a mobile station was assigned to each quadrant to capture the number of mobile users in each quadrant by aggregating it for different time periods (Figure 4).

For the analysis of human commuting, tree-type directed graphs were built, and the commuting was analysed between nodes on edges, which show commuting between certain quadrants (representing the territory of a municipality). To calculate the results, a statistical ensemble model was built. The model consists of several parts: a Wasserstein distance calculation statistical model, a graph data structure and graph model for the calculation of the graph direction and nodes values (see Appendix A).

**FIGURE 4** The division of the territory of Latvia into regular squares of $5 \times 5$ km$^2$ with mobile stations
To calculate the commuting of mobile phone users across the directed graph, the following algorithm was applied to each $5 \times 5$ km$^2$, where standard Python libraries and a Wasserstein distance calculation were used for data preparation, data load, and Python library import:

```python
# lib to calculate our edge lengths
from scipy.stats import wasserstein_distance
# lib that we'll use to build the graph
import networkx as nx
# lib for drawing our graph
import matplotlib.pyplot as plt
# lib for loading csv files easily
import csv
# lib for using regular expressions (we use it to process the csv rows)
import re
# data structures to hold our csv data
mobile_notikumi = {}
kvadrati = {}
kvadrati_un_stacijas = {}
```

Before running the analysis, three data sets had to be prepared: a matrix of quadrants, mobile activity, and mapping between quadrants and mobile stations. It is vitally important to choose the right territory for analysis due to the algorithm's sensitivity to the chosen quadrant size matrix.

For the territory of Latvia, the 5-km quadrant step is relatively large. However, due to the low density of population (on average 30 inhabitants per 1 km$^2$) in some parts of the territory, municipality quadrants matrix had to be expanded up to 20-km quadrants, meaning that a larger territory would be analysed in terms of commuting in and out of the territory.

The quadrant matrix was prepared as a CSV (comma-separated values) file and had to be fully filled. If the analysed territory is not as square as all other areas, it should be marked with a zero. Then the data which mapped the activity between quadrants and mobile stations was uploaded.

Afterwards, the hourly data of mobile activity for each mobile station for 1 day was uploaded. The data could be prepared for a specific day or as an average of a period, for all workdays for a month or a year etc. This allows an analysis of commuting patterns for different periods, which represent a day's average. Using the count of active unique users per mobile station within a specific time of the day enables a more vivid prediction of human commuting than using total call activities per mobile station, because many activities may be performed by the same user.

When data is loaded using grid and mobile stations mapping information, a model was created to produce a directed graph to calculate the number of unique mobile phone users on each node and calculate the direction on edges. In this case, a tree-type graph was used to avoid circular dependencies and calculations. The graph, built out of the loaded data structures, was converted into matrix.

The graph indicates direction and changes between nodes found by calculating hourly changes on the nodes. This was done in loops by calculating hourly changes and assigning the final number of unique mobile users for each node. The Python code for calculating the directions on the edges is shown in Appendix A. The results of the analysis were generated in a CSV file with information on commuting between quadrants and the direction and the number of commuters in a particular hour.

## RESULTS

Based on the PCA-obtained results, the 110 municipalities and 9 cities of Latvia were divided into eight groups (Table 1) according to their mobile phone activity for three time periods: 2015–2016, 2017, and 2018 (taking into account economic activity on workdays and weekends/holidays). The efficiency of the economic activity strategy was calculated, using a Formula (1), and compared with 40% to 100% efficiency curves. The efficiency curve depicts the level to which each municipality uses its potential and allows a comparison of the performance of different regions.

### 3.1 Dividing municipalities into groups

The division of the 110 municipalities and 9 cities of Latvia into groups according to their mobile phone activity for 2015–2016 and 2018 is shown accordingly in Figures 5 and 6.
The first group, “Hard Workers,” is characterized by high activity on workdays but low activity on average on weekends/holidays. This group is the driving force behind the economy of Latvia but does not fully exploit its holiday potential. These municipalities are highly dependent on fluctuations in economic activity; therefore, the service sector should be developed. For example, Riga, the capital city of Latvia, is the central metropolis of the Baltic countries and an international infrastructure hub (The Freeport of Riga, Riga International Airport).

The city of Riga is characterized by transit and logistics companies and a developing social infrastructure. At the same time, it is a monocentric city, which insufficiently uses its tourism potential, with only 5% of small and medium enterprises (SMEs) operating in this sector. Riga has congested traffic infrastructure and a shrinking population.

The city of Jelgava is located only 40 km away from Riga and has excellent infrastructure. It is also home of the Latvia University of Life Sciences and Technologies. However, at weekends and on holidays, business activity is low there, and leisure opportunities are insufficient, as only 4% of SMEs operate in the tourism sector.

Both Riga and Jelgava fall into the group of “Hard Workers,” with high activity on workdays but low activity on average on weekends/holidays, as well as a strong negative seasonal effect during the summer. If these cities do not think about sustainable development, in the long term, their populations may decrease.

The second “Congruent” group of Latvia’s municipalities is characterized by high and moderate activity on workdays and average activity on weekends/holidays. The group has balanced development but insufficient resources for the next breakthrough. Depending on the priorities, it is necessary to develop the production or the service sector. The city of Ventspils has a developed production and logistics sector, a well-developed port, is home to Ventspils University College, and 5% of its SMEs operate in the tourism sector.

The “Holidaymakers” group is characterized by low activity on average on workdays but high activity on average on weekends/holidays. These municipalities make sufficient use of their leisure potential but not the business one. It is necessary to develop the production sector and change the municipal development strategy.

For example, Jurmala is a historic sea resort near Riga, with a favourable strategic location; today, 7% of SMEs there operate in the tourism sector. However, it does not fully use its holiday potential during the off-season. The development of Jurmala is too dependent on the solvency of the population, as during economic recessions the demand for leisure and entertainment services tends to decrease.
Whereas the city of Ventspils, in the "Congruent" group, is characterized by high and moderate activity on workdays and average activity on weekends/holidays and during summer time, Jurmala, from the "Holidaymakers" group, has low activity on average on workdays but high activity on average on weekends/holidays, as well as a strong positive seasonal effect during the summer time. The average number of mobile phone call activities on LMT per day in Riga, Jelgava, Ventspils, and Jurmala by month in 2018 is shown in Figure 7.
The “Moderate” group is characterized by average economic activity on all days, meaning that their resource potential is not being used sufficiently. It is necessary to increase labour productivity and economic potential; otherwise, economic activity and regional development will decrease.

The “Disinterested” group is characterized by low activity on workdays and average activity on weekends/holidays. There is holiday potential but low economic activity on workdays. It is necessary to develop the service sector and to change the development strategy of the region in order to avoid degradation. For example, in Cibla municipality, the overall economic activity is very low. Only 1% of SMEs, or three companies, were active in the tourism sector.

The “Party Makers” group is characterized by low activity on workdays but high activity on weekends/holidays. It is necessary to develop the production sector and change the region’s development strategy. There is a high dependence on the purchasing power of the population. For example, Salacgriva municipality is located by the sea and fully exploits its potential as a holiday destination during the summer months, but on workdays and during the off-season, its economic activity is minimal. Ten percent of SMEs there operate in the tourism sector.

The “Hedonists” group is characterized by the lowest activity on workdays but the highest activity on holidays. There is no economic potential for the manufacturing sector. It is necessary to develop the production sector and change the region’s development strategy. There is maximum dependence on the purchasing power of the population. For example, Rucava municipality has a strong positive seasonal effect in the summer, making it unique—the number of mobile phone activities during holidays is higher there than on workdays (Figure 8).

The “Phenomenon” group is characterized by average activity on workdays and moderately low activity on weekends/holidays: in 2015–2016, the only member of this group was Vilani municipality; in 2017, Rugaji municipality; but in 2018, Ergli municipality.

3.2 The economic activity efficiency of Latvia’s municipalities

The grouping of the municipalities by economic activity is shown (Figure 9), using the principal components loads with the data aggregated by days. Latvia’s municipalities are divided into eight groups according to their economic activity.

The negative second PC values in the “Hedonists” group means that the mobile phone activity during holidays, in absolute values, is higher than on workdays. It is characteristic only of the “Hedonists” group: Saulkrasti municipality in 2015–2016 and Rucava municipality in 2015–2016, 2017, and 2018.

The groups provide an understanding of the profile of each municipality, but the efficiency curve makes it possible to assess the efficiency of the strategy chosen by each municipality (Figure 9).

When compared with the 2015–2016 period, in 2017, the economic activity in Latvia’s municipalities improved, as is indicated by the distribution of groups. For example, the number of municipalities included in the “Disinterested” group decreased twice, and the number of municipalities between 95% and 100% efficiency curves doubled in 2018 in comparison to the previous year.

The increase of economic activity in the regions of Latvia relates to the country’s overall economic growth and corresponds to the activities implemented by the regions. The results of the analysis allow us to compare the development dynamics of different regions and to evaluate their sustainability.

Therefore, it is important to continue to monitor the economic activity in the municipalities in order to assess the sustainability of their chosen development strategies (Figure 6). The distribution of all of Latvia’s 119 municipalities according to the efficiency criterion EC shows that the number of municipalities lies between two different efficiency curves (Figure 9):

- From 95% to 100%: 16 municipalities in 2017 and 33 municipalities in 2018.
- From 90% to 95%: 31 municipalities in 2017 and 29 municipalities in 2018.
- From 80% to 90%: 36 municipalities in 2017 and 32 municipalities in 2018.
- From 40% to 80%: 36 municipalities in 2017 and 25 municipalities in 2018.

![Figure 8](image-url) The average number of mobile phone call activities per day in Rucava municipality by month and workdays, 2018
3.3 | The distribution of regions by human commuting patterns

A visual graph representation for each hour shows the complexity and the variety of the commuting patterns within the analysed territory. It also enables us to visualize the central nodes of the graph. Interim results of analysis available as a graph are shown in Figure 10.

An analysis of large territories represented as graphs does not give the value needed, and other tools are required to interpret the results of this research, such as GIS software (e.g., Quantum GIS, used in this research). For the analysis in GIS systems, the results are generated in a tabular form (Figure 11). The results (in a tabular form) are exported as a CSV file containing the following columns:
• [Num]: row number, used as a record code.
• [Source]: quadrant number where commuting starts (from quadrant).
• [x_f]: longitude of source quadrant centroid.
• [y_f]: latitude of source quadrant centroid.
• Target: quadrant number where commuting ends (destination quadrant).
• [x_t]: longitude of destination quadrant centroid.
• [y_t]: latitude of destination quadrant centroid.
• Score: number of unique users commuting.

The resulting table contains all discovered commutes, where each row represents a commute from quadrant to quadrant. This allows us to use these results for a visual representation on a map as arrows from source quadrant to target (destination) quadrant. For the visualization of the results, Open Source Geographic Information Software, QGIS, was used. For a full representation of all the layers needed for the interpretation of the results, the following data layers were used to represent data in a graphic form:

- A geographic map of Latvia—shape file.
- Regular 5 × 5 km²—shape file.
- Latvia’s municipalities—shape file.

On the top of the geographic layers, the calculated commuting data layers (represented as arrows) were placed. Two calculated layers of arrows were created, one for the time period between 7 a.m. and 9 a.m. (as black arrows); and the second for the time period between 5 p.m. and 7 p.m. (as white arrows). For the assessment of the efficiency and self-sufficiency of municipality centres, the Top 75 general commuting routes (and direction) were calculated for the entire territory of Latvia, which allows us to understand general commuting patterns throughout the country. This could also be used as a general overview to understand territories with large activities.

Due to the sensitivity of the applied algorithm, more localized territories with the volume of mobile activity in absolute numbers must be used for the analysis of each municipality. Mobile activity in Latvia is unevenly distributed across the country, with concentrations in a couple of larger cities; just like the density of mobile stations per quadrant in large cities and their nearby areas can be over 10 times higher than in other territories. See the average commuting directions during workdays in October for 2017 and 2018 in Figures 12 and 13.

**FIGURE 12** The Top 75 human commuting directions in Latvia from 7 a.m. to 9 a.m. on workdays
The analysis of the Top 75 commuting routes during morning (Figure 12) and evening hours (Figure 13) shows outgoing and incoming patterns. For some territories, black arrows, which represent the morning commutes, show movement to larger municipal centres and territories with a higher economic and social activity. In the evening, commuters move into the opposite direction—to their places of residence.

For example, Riga attracts economically active people who live in nearby municipalities. There is a clear commuting pattern of people from nearby municipalities who come to Riga during the morning hours—Riga is attractive, with its better-paid jobs as well as its social, educational, and cultural infrastructure, such as schools, theatres, opera, etc. (Figure 14).

Because the regions of Latvia differ in terms of self-sufficiency, the information on the commuting patterns can be used as an additional factor in evaluating regional development.

Regions with high economic activity are also more self-sufficient, and this can be measured using commuting information. The latter can be used as an indicative value because it is influenced by other factors like transit, which should be analysed separately.
4 | CONCLUSIONS

The volume of mobile phone call activities is an indicator of economic activity in municipalities and larger regions and can be used in developing a reliable tool for continuous and dynamic monitoring of the region's economic performance.

There is a seasonal effect on mobile phone activities, as well a significant difference between workdays and weekends/holidays, which distinguishes municipalities and regions with different economic activity patterns.

The data obtained from Latvia's municipalities resulted in the identification of eight distinct groups representing unique patterns of economic activity. The efficiency curve allows us to evaluate the effectiveness of the development strategy chosen by each municipality. The characteristics of these groups depend on economic activity on workdays and weekends/holidays, as well as the seasons.

In 2017, the economic activity in Latvia's municipalities improved compared with the period of 2015–2016, as indicated by the distribution in the groups. For example, the number of municipalities included in the "Disinterested" group decreased by half. In 2018, the number of municipalities between the 95% and 100% efficiency curves doubled in comparison to 2017.

The economic centres among Latvia's municipalities have high activity on weekdays and moderate or medium activity on weekends/holidays. They are the driving force of the Latvian economy but do not fully exploit their holiday potential. Depending on the priorities, the manufacturing or the service sector could be developed; wrongly chosen priorities might fragment the available resources.

The service centres in the regions of Latvia are characterized by low activity on weekdays and moderate activity at weekends—a good use of holiday potential but not of the workday potential. Changes to a region's development strategy should be considered, as there is a moderate dependence on the purchasing power of the population.

The authors propose the regional development index as a real-time or periodic monitoring tool. The method developed here provides a practical tool for regional governments to keep track of their strategy implementation and enables a strategic gap analysis.

This method also provides a dynamic visualization of the strategic direction a particular municipality has achieved between the periods of measurement and can be used as an additional central performance indicator if regularly measured by regional governments. The developed method was tested on Latvian mobile telephone data sets as proxies, and the regional development index was created.

The regional economic activity efficiency evaluation is based on mobile phone data statistics, taking into account the requirements of the GDPR, where data about individual calls switching was not accessed, and personal data protection was respected.

In addition to the regional economic activity index, human commuting patterns can be used to show the approximate directions of people's movements across municipal territories. Human commuting activities indicate the attractiveness and economic self-sufficiency of a municipality or other territory.

The internal mobility of the population depends on the day of the week, as evidenced by the daily changes in the intensity of call events. Two intervals, from 7 a.m. to 9 a.m. and from 5 p.m. to 7 p.m. in 2017 and 2018, were selected to analyse fluctuations in the intrinsic activity of the population on a weekly basis.

The study concluded that on a workday, the movement of the population is more intense, compared with weekends and holidays. From 7 a.m. to 9 a.m., the flow is in the direction of economic centres, but from 5 p.m. to 7 p.m., in the opposite direction.

The collected data and the developed method provided a reliable tool to evaluate and visualize the internal mobility of the population and its dependency on the month of the year, as evidenced by the change in the intensity of call events in different months of the year.

ACKNOWLEDGEMENTS

The University of Latvia and LMT Ltd. (Grant AAP2016/B089) supported this work. The University of Latvia and the Ministry of Environmental Protection and Regional Development of the Republic of Latvia (Grant IL/25/2019) supported this work.

ORCID

Irina Arhipova  https://orcid.org/0000-0003-1036-2024

REFERENCES

Ahas, R., Aasa, A., Mark, U., Pae, T., & Kull, A. (2007). Seasonal tourism spaces in Estonia: Case study with mobile positioning data. Tourism Management, 28, 898–910.

Ahas, R., Aasa, A., Roose, A., Mark, U., & Silm, S. (2008). Evaluating passive mobile positioning data for tourism surveys: An Estonian case study. Tourism Management, 29, 469–486.

Ahas, R., Aasa, A., Silm, S., & Tiru, M. (2010). Daily rhythms of suburban commuters' movements in the Tallinn metropolitan area: Case study with mobile positioning data. Transportation Research Part C, 18(1), 45–54.

Alexander, L., Jiang, S., Murga, M., & Gonzalez, M. C. (2015). Origin—Destination trips by purpose and time of day inferred from mobile phone data. Transportation Research Part C, 58, 240–250.

Arhipova, I., Berzins, G., Brekis, E., Binde, J., Opmanis, M., 2019. Mobile phone data statistics as proxy indicator for regional economic activity assessment. Proceedings of the 1st International Conference on Finance, Economics, Management and IT Business (FEMIB 2019), Crete, Greece, pp. 27–36.
Arhipova, I., Berzins, G., Brekis, E., Kravcova, J., Binde, J. (2017). The methodology of region economic development evaluation using mobile positioning data. In D. Cingula, ed. 20th International Scientific Conference on Economic and Social Development. Prague: Varazdin Development and Entrepreneurship Agency, Varazdin, Croatia. City of Prague. University North, Koprivnica, Croatia. Faculty of Management University of Warsaw, Warsaw, Warsaw, Poland, pp. 111–120.

Arhipova, I., Berzins, G., Brekis, E., Opmanis, M., Binde, J., Steinbuka, I., & Kravcova, J. (2019). Pattern identification by factor analysis for regions with similar economic activity based on mobile communication data. Adv. Intell. Syst. Comput., 886, 561–569.

Balzotti, C., Bragagnini, A., Brianii, M., & Cristiani, E. (2018). Understanding human mobility flows from aggregated mobile phone data. IFAC-PapersOnLine, 51(9), 25–30.

Bonnel, P., Hombourger, E., Olteau-Raimond, A.-M., & Smoreda, Z. (2015). Passive mobile phone dataset to construct origin-destination matrix: Potentials and limitations. Transportation Research Procedia, 11, 381–398.

Bwambale, A., Choudhury, C. F., & Hess, S. (2019). Modelling trip generation using mobile phone data: A latent demographics approach. Journal of Transport Geography, 76, 276–286.

Chang, V. (2017). Towards data analysis for weather cloud computing. Knowledge-Based Systems, 127, 29–45.

Chang, V. (2018). An overview, examples, and impacts offered by emerging services and analytics in cloud computing virtual reality. Neural Computing and Applications, 29(5), 1243–1256.

Chen, C., Bian, L., & Ma, J. (2014). From traces to trajectories: How well can we guess activity locations from mobile phone traces? Transportation Research Part C, 46, 326–337.

Chetty, G., White, M., & Akther, F. (2015). Smart phone based data mining for human activity recognition. Procedia Computer Science, 46, 1181–1187.

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F. R., Gaughan, A. E., ... Tatem, A. J. (2014). Dynamic population mapping using mobile phone data. PNAS (Proceedings of the National Academy of Sciences of the United States of America), 111(45), 15888–15893.

Hongyan, G., & Fasheng, L. (2013). Estimating freeway traffic measures from mobile phone location data. Transportation Research Procedia, 5, 278–291.

Elías, D., Nadler, F., Stehno, J., Krosche, J., & Lindorfer, M. (2016). SOMOBIL—Improving public transport planning through mobile phone data analysis. Transportation Research Procedia, 14, 4478–4485.

European Parliament. (2016). Regulation (EU) 2016/679 of the European parliament and of the council of April 27, 2016 on the protection of personal data. Official Journal of the European Union, L119, 1–88.

Hatzioannidou, F., & Polydoropoulou, A. (2017). Passenger demand and patterns of tourists’ mobility in the Aegean archipelago with combined use of big datasets from mobile phones and statistical data from ports and airports. Transportation Research Procedia, 25, 2309–2329.

Lee, Y.-S., & Cho, S.-B. (2014). Activity recognition with android phone using mixture-of-experts co-trained with labeled and unlabeled data. Expert Systems with Applications, 41(14), 6174–6189.

Larijani, A. N., Olteanu-Raimond, A.-M., Perret, J., Bredif, M., & Ziemlicki, C. (2015). Investigating the mobile phone data to estimate the origin destination flows and analysis: case study: Paris region. Transportation Research Procedia, 6, 64–78.

Lee, J., Hong, B., Jung, S., & Chang, V. (2018). Clustering learning model of CCTV image pattern for producing road hazard meteorological information. Future Generation Computer Systems, 86, 1338–1350.

Lee, K.-S., You, S. Y., Eom, J. K., Song, J., & Min, J. H. (2018). Urban spatiotemporal analysis using mobile phone data: Case study of medium- and large-sized Korean cities. Habitat International, 73(October 2017), 6–15.

Lee, Y.-S., & Cho, S.-B. (2014). Activity recognition with android phone using mixture-of-experts co-trained with labeled and unlabeled data. Neurocomputing, 126, 106–115.

Liu, F., Janssens, D., Cui, J. X., Wang, Y. P., Wets, G., & Cools, M. (2014). Building a validation measure for activity-based transportation models based on mobile phone data. Expert Systems with Applications, 41(14), 6174–6189.

Liu, F., Janssens, D., Wets, G., & Cools, M. (2013). Annotating mobile phone location data with activity purposes using machine learning algorithms. Expert Systems with Applications, 40(8), 3299–3311.

Ni, L., Wang, X. C., & Chen, X. M. (2018). A spatial econometric model for travel flow analysis and real-world applications with massive mobile phone data. Transportation Research Part C, 86, 510–526.

Oliveira, E. M. R., Viana, A. C., Naveen, K. P., & Sarraute, C. (2017). Mobile data traffic modeling: Revealing temporal facets. Computer Networks, 112, 176–193.

Phithakkitnukoon, S., Horanont, T., Wityayangkum, A., Siri, R., Sekimoto, Y., & Shibasaki, R. (2015). Understanding tourist behavior using large-scale mobile sensing approach: A case study of mobile phone users in Japan. Pervasive and Mobile Computing, 18, 18–39.

Raun, J., Ahas, R., & Turu, M. (2016). Measuring tourism destinations using mobile tracking data. Tourism Management, 57, 202–212.

Rios, S. A., & Muñoz, R. (2017). Land use detection with cell phone data using topic models: Case Santiago, Chile. Computers, Environment and Urban Systems, 61, 39–48.

Sørensen, A. Ø., Bjelland, J., Bull-Berg, H., Landmark, A. D., Akhtar, M. M., & Olsson, N. O. E. (2018). Use of mobile phone data for analysis of number of train travellers. Journal of Rail Transport Planning & Management, 8(2), 123–144.
Steinbruggen, J., Tranos, E., & Nijkamp, P. (2015). Data from mobile phone operators: A tool for smarter cities? *Telecommunications Policy*, 39(3–4), 335–346.

Steinbruggen, J., Tranos, E., & Rietveld, P. (2016). Traffic incidents in motorways: An empirical proposal for incident detection using data from mobile phone operators. *Journal of Transport Geography*, 54, 81–90.

Sun, G., Chang, V., Yang, G., & Liao, D. (2018). The cost-efficient deployment of replica servers in virtual content distribution networks for data fusion. *Information Sciences*, 432, 495–515.

Sun, S., Xiong, C., & Chang, V. (2019). Acceptance of information and communication technologies in education: An investigation into university students’ intentions to use mobile educational apps. *International Journal of Enterprise Information Systems*, 15(1), 24–44.

Trasarti, R., Olteanu-Raimond, A.-M., Nanni, M., Couronne, T., Furletti, B., Giannotti, F., … Ziemlicki, C. (2015). Discovering urban and country dynamics from mobile phone data with spatial correlation patterns. *Telecommunications Policy*, 39(3–4), 347–362.

Wang, Y., Correia, G. H. d. A., Arem, V. B., & Timmermans, H. J. P. (2018). Understanding travellers’ preferences for different types of trip destination based on mobile internet usage data. *Transportation Research Part C*, 90, 247–259.

AUTHOR BIOGRAPHIES

Irina Arhipova is the professor of Econometrics at the Department of Management Systems of the Faculty of Information Technologies at Latvia University of Life Sciences and Technologies (LLU), as well as the lead researcher at the University of Latvia. Irina Arhipova is one of Latvia’s most experienced researchers in Econometrics and Experimental Economics, the author of over 100 scientific articles, a frequent speaker at international scientific conferences and a guest lecturer at foreign universities.

Gundars Berzins is the professor and the Dean of the Faculty of Business, Management and Economics at the University of Latvia. He holds a doctorate in management science and as a leading researcher has participated in many research projects and has authored a number of international scientific publications. His research interests include business strategies and regional policy. He has previously held the post of Chancellor of the University of Latvia. His passion is the importance of strategy and strategic thinking in business. He is excited at reading lectures on “Management theories”, “Business Strategy and Policy Management” and “Strategic and Operational Planning”.

Edgars Brekis is the Professor at the University of Latvia and one of the best data analysts in Latvia, a lecturer and the author of scientific publications. His latest research: Big Data Processing Mathematical Model for the Updatable Latvian Regional Business Index, using limited data asset; Expert Evaluation on the Improvement of Macroeconomic Indicator Forecasts and the Improvement of MediumTerm Macroeconomics Forecast Model.

Juris Binde is the CEO of Latvia’s leading mobile telecommunications operator Latvia Mobile Telephone (LMT), and a professor at Vidzeme University of Applied Sciences (VIA). On a regular basis, Juris Binde speaks at scientific and professional conferences, is the author of tens of research publications, a number of which have been published in internationally reviewed journals on management science and business management, information technologies, data processing and the Internet of things.

Martins Opmanis is researcher at the Institute of Mathematics and Computer Science of University of Latvia. He is team leader of Latvian team at International Olympiads in Informatics for high-school students since 1996. His main scientific interests are processing of big data, correctness of data analysis, organizing and managing various competitions in Informatics and Mathematics.

Aldis Erglis is a PhD student in Management Studies Programme at the University of Latvia, Faculty of Business, Management and Economics. The topic of his Doctoral thesis is Dynamic Management of the Organization – Linking Strategy to Operations. Aldis Erglis has a unique experience in complex IT solution development and programming, data analysis and visualisation.

Evija Ansonska has obtained a MA degree in Political Science at the University of Latvia, has studied political science in Denmark. Currently, she is the Head of Marketing and Communication at the University of Latvia, Faculty of Business, Management and Economics and a research assistant. Evija Ansonska has a 17-year experience in communication and project management. As a project manager and a research assistant, Ansonska has participated in different research projects.
APPENDIX A.

Python code for edge direction calculation:

Calculating the minimum spanning tree and figuring out what the delta for each hour is
```
```python
# get the minimum spanning tree as it's going to be the data struct that we'll be basing our calculations on
t = nx.minimum_spanning_tree(g)
# translate the distribution into deltas so that we can see the change from the past hour
notikumu_pieaugums = {}
for key, values in kvadratu_notikumi.items():
    notikumu_pieaugums[key] = [(values[i]-values[(i-1)%len(values))] for i in range(len(values))]
# we need to transform the data in a way that let's work with it in a hour by hour basis instead of a square by square basis
stundu_notikumu_pieaugums = {}
for key, values in notikumu_pieaugums.items():
    # go through all of the hours
    for i in range(len(values)):
        # and add this squares score in this hour to the new data structure
        if i not in stundu_notikumu_pieaugums:
            stundu_notikumu_pieaugums[i] = {}
            stundu_notikumu_pieaugums[i][key] = values[i]
```
Calculating the minimum spanning tree and figuring out what the delta for each hour is
```
```python
# get the minimum spanning tree as it's going to be the data struct that we'll be basing our calculations on
t = nx.minimum_spanning_tree(g)
# translate the distribution into deltas so that we can see the change from the past hour
notikumu_pieaugums = {}
for key, values in kvadratu_notikumi.items():
    notikumu_pieaugums[key] = [(values[i]-values[(i-1)%len(values))] for i in range(len(values))]
# we need to transform the data in a way that let's work with it in a hour by hour basis instead of a square by square basis
stundu_notikumu_pieaugums = {}
for key, values in notikumu_pieaugums.items():
    # go through all of the hours
    for i in range(len(values)):
        # and add this squares score in this hour to the new data structure
        if i not in stundu_notikumu_pieaugums:
            stundu_notikumu_pieaugums[i] = {}
            stundu_notikumu_pieaugums[i][key] = values[i]
```
Figuring out the directions of the edges
```
```python
# let's make a graph visual for each hour
# without updating any of the scores
hour_graphs_static = {}
# with updating the scores
hour_graphs_updating = {}

for h, sq in stundu_notikumu_pieaugums.items():
    print(h)
    print(sq)
    # make the current hours graph
    hour_graphs_static[h] = nx.DiGraph()
    hour_graphs_updating[h] = nx.DiGraph()
    # let's deal with the non-updating graph first
    marked = []
    # just want to cover all of the nodes
    while True:
        # get a list of values that we haven't used as center nodes yet
        workable = [v for k, v in sq.items() if k not in marked]
        # only continue if there's at least 2 more nodes to work with
        if len(workable) < 2:
            break
        min_v = min(workable)
        for k, v in sq.items():
            if v == min_v and k not in marked:
                # found the key of the minimum node, now to update the graph accordingly
                marked.append(k)
                # find how many edged are associated with this node
                edge_count = len([e for e in t.edges() if e[0] == k or e[1] == k])
        for e in t.edges():
            if e[0] == k and not hour_graphs_static[h].has_edge(k, e[1]) and not hour_graphs_static[h].has_edge(e[1], k):
                # add an edge to our directed graph
                hour_graphs_static[h].add_edge(k, e[1], weight=int((sq[e[1]] - sq[e[0]])/edge_count))
            if e[1] == k and not hour_graphs_static[h].has_edge(k, e[0]) and not hour_graphs_static[h].has_edge(e[0], k):
                # add an edge to our directed graph
                hour_graphs_static[h].add_edge(k, e[0], weight=int((sq[e[0]] - sq[e[1]])/edge_count))
        break
    # let's deal with the updating dataset now
    marked = []
    # we're going to be manipulating this
    tmp_sq = sq.copy()
    # just want to cover all of the nodes
    while True:
        # get a list of values that we haven't used as center nodes yet
        workable = [v for k, v in tmp_sq.items() if k not in marked]
        # only continue if there's at least 2 more nodes to work with
        if len(workable) < 2:
            break
        min_v = min(workable)
        for k, v in tmp_sq.items():
            if v == min_v and k not in marked:
                # found the key of the minimum node, now to update the graph accordingly
                marked.append(k)
                # find how many edged are associated with this node
                edge_count = len([e for e in t.edges() if e[0] == k or e[1] == k])
        for e in t.edges():
            if e[0] == k and not hour_graphs_static[h].has_edge(k, e[1]) and not hour_graphs_static[h].has_edge(e[1], k):
                # add an edge to our directed graph
                hour_graphs_static[h].add_edge(k, e[1], weight=int((sq[e[1]] - sq[e[0]])/edge_count))
            if e[1] == k and not hour_graphs_static[h].has_edge(k, e[0]) and not hour_graphs_static[h].has_edge(e[0], k):
                # add an edge to our directed graph
                hour_graphs_static[h].add_edge(k, e[0], weight=int((sq[e[0]] - sq[e[1]])/edge_count))
        break
if e[0] == k and not hour_graphs_updating[h].has_edge(k, e[1]) and not hour_graphs_updating[h].has_edge(e[1], k):
    # add an edge to our directed graph
    hour_graphs_updating[h].add_edge(k, e[1], weight=int((tmp_sq[e[1]] - tmp_sq[e[0]]) / edge_count))
    # and update the scores to reflect
    print("looking at " + str(k) + " and updating " + str(e[1]) + " from " + str(tmp_sq[e[1]]) + " to " + str(tmp_sq[e[1]] + int((tmp_sq[e[1]] - tmp_sq[e[0]]) / edge_count)))
    tmp_sq[e[1]] += int((tmp_sq[e[1]] - tmp_sq[e[0]]) / edge_count)
if e[1] == k and not hour_graphs_updating[h].has_edge(k, e[0]) and not hour_graphs_updating[h].has_edge(e[0], k):
    # add an edge to our directed graph
    hour_graphs_updating[h].add_edge(k, e[0], weight=int((tmp_sq[e[0]] - tmp_sq[e[1]]) / edge_count))
    # and update the scores to reflect
    print("looking at " + str(k) + " and updating " + str(e[0]) + " from " + str(tmp_sq[e[0]]) + " to " + str(tmp_sq[e[0]] + int((tmp_sq[e[0]] - tmp_sq[e[1]]) / edge_count)))
    tmp_sq[e[0]] += int((tmp_sq[e[0]] - tmp_sq[e[1]]) / edge_count)
break
print(tmp_sq)
```