Estimating Behaviour Patterns and Activity Ranges of Minors Across Japan for Effective Infectious Disease Control

Takahiro Yoshimi[1], Yuki Akiyama[2], Ryosuke Shibasaki[3], Zekun Cai[4], Yiheng Yao[5]

[1]Tokyo City University, Setagaya-ku, Tokyo, Japan – g2481652@tcu.ac.jp
[2]Tokyo City University, Setagaya-ku, Tokyo, Japan – akiyamay@tcu.ac.jp
[3]LocationMind, inc., Chiyoda-ku, Tokyo, Japan – ryosuke.shibasaki@gmail.com
[4]LocationMind, inc., Chiyoda-ku, Tokyo, Japan – ze-cai@locationmind.com
[5]The University of Tokyo, Bunkyo-ku, Tokyo, Japan – 7665643375@g.ecc.u-tokyo.ac.jp

Keywords: Minors, Human Mobility Data, Population Distribution, Infectious Disease Control, Statistics.

Abstract

The COVID-19 pandemic has heightened public awareness of infectious disease prevention. However, minors often lack fully developed hygiene awareness, making it difficult to enforce basic preventive measures such as handwashing and mask-wearing. In addition, their frequent close contact in enclosed environments, such as educational institutions, increases the risk of rapid disease transmission. As minors can serve as significant vectors of infection, understanding their movement patterns is essential for devising effective countermeasures. Despite this need, tracking minors using mobile device location data remains difficult due to the low smartphone ownership rate. Therefore, as a first step toward overcoming the limitations of mobile device tracking and enabling large-scale behavioural analysis of minors, this study developed a method for estimating their residential locations at a micro-scale, such as the building level, using various statistical datasets. Furthermore, it estimated their daily commuting destinations—such as schools and kindergartens—as well as their daily schedules and commuting routes, in order to project their routine behavioural patterns and spatial activity ranges on a nationwide scale. This approach enables a more comprehensive understanding of minors' daily movements and supports the planning and implementation of more effective infectious disease prevention strategies.

1. Introduction

The novel coronavirus disease (COVID-19), first identified in December 2019 in Wuhan, Hubei Province, China, rapidly spread across the globe, resulting in extensive worldwide impact. According to data compiled by Johns Hopkins University, the United States reported the highest cumulative number of confirmed cases, followed by India, France, Germany, Brazil, and Japan. Specifically, the United States recorded approximately 103.8 million infections and 1.12 million deaths, while Japan reported around 33.3 million infections and 73,046 deaths (Center for Systems Science and Engineering (CSSE), 2023).

In response to the profound impact of the COVID-19 pandemic, emergency declarations and targeted measures to prevent the spread of infectious diseases were implemented multiple times in Japan. These actions contributed to increased public awareness of infection control measures.

According to the Basic Knowledge of Infection Control Measures published by Japan's Ministry of Health, Labor and Welfare, infectious diseases arise through the interaction of three key factors: the pathogen (infection source), the route of infection, and the host. Effective infection control strategies, therefore, depend on disrupting at least one of these factors, with particular emphasis on blocking transmission routes (Ministry of Health, Labour and Welfare, n.d.).

There are four primary routes of transmission: airborne, droplet, contact, and oral. As outlined in the Guidelines for the Prevention of Infectious Diseases in Nursery Schools, issued by the Administration for Ministry of Health, Labour and Welfare (MHLW), it is essential to implement appropriate countermeasures tailored to each transmission route (Ministry of Health, Labour and Welfare (MHLW), 2021).

Airborne transmission occurs when small respiratory droplets expelled by an infected person during coughing, sneezing, or speaking dry out, leaving behind droplet nuclei that contain the pathogen. These nuclei can remain suspended in the air, retain their infectivity, and be inhaled by others. Since airborne transmission often occurs in enclosed spaces, ensuring proper ventilation is a critical preventive measure.

Droplet transmission occurs when an infected individual coughs, sneezes, or talks, releasing droplets that contain pathogens. These droplets can be inhaled by individuals in close proximity. Preventive strategies focus on minimizing exposure to these droplets by avoiding crowded places and close-contact settings where conversations or vocalizations occur. In addition, strict adherence to cough etiquette and the use of face masks are essential.

Contact transmission can occur via two main pathways: direct contact, such as shaking hands, holding, or kissing an infected person; and indirect contact through contaminated surfaces or objects, including doorknobs, handrails, or playground equipment. Typically, infection does not occur simply by having pathogens on the surface of the body—it is established only when the pathogen enters the body. Preventive measures include maintaining good hand hygiene through regular handwashing, gargling, and the use of hand sanitizers.

Oral transmission occurs when pathogens are ingested through contaminated food or water, eventually reaching the gastrointestinal tract and causing infection. Preventing oral transmission requires proper hygiene management during food preparation and handling, following relevant guidelines and official recommendations.

Nevertheless, minors—particularly young children—often lack sufficient awareness of hygiene practices, making it difficult for them to consistently adhere to basic infection prevention measures such as handwashing and mask-wearing. In addition, they frequently spend time in densely populated environments such as schools and childcare facilities, and their immune systems are still underdeveloped. As a result, they face a higher risk of rapidly transmitting infections to their families and surrounding communities, potentially becoming hotspots for the spread of infectious diseases.

1.1 Utilization of Human Mobility Data in Infectious Disease Control

Human mobility data can serve as a critical resource in formulating effective infectious disease control strategies. By analysing such data, researchers and policymakers can gain a deeper understanding of how infectious diseases spread, allowing them to identify high-risk areas and time periods with greater precision. For instance, Doi and Onishi (2023) evaluated the impact of Japan's state of emergency declarations and their subsequent lifting during the COVID-19 pandemic by utilizing long-term, nationwide mobility data derived from mobile device location information. Similarly, Kawakami et al. (2023) examined the relationship between COVID-19 case numbers and human mobility in Japan, demonstrating that mobility restrictions—such as those enacted during emergency declarations—were effective in curbing the spread of the virus. In another study, Mishima et al. (2022) conducted experiments using two types of mobility data—mesh-based population distribution data and inter-mesh population transition data—to predict the number of confirmed COVID-19 cases during the third wave in Tokyo between 2020 and 2021. Their findings emphasized the significance of population transition information in modelling human-mediated phenomena, including infectious disease transmission. Furthermore, Onoda et al. (2020) emphasized the necessity of precisely monitoring and adapting to rapidly changing human mobility patterns in the context of the "with-COVID" and "post-COVID" eras. They advocated for the potential application of Mobile Spatial Statistics, which are estimated using operational data from approximately 80 million mobile phones, as a means of achieving this goal.

Comparable studies have been conducted in countries outside Japan. Ilin et al. (2021) demonstrated that publicly available mobility data from platforms such as Google, Facebook, Baidu, and SafeGraph can be effectively used to assess the impact of mobility restriction policies and to forecast the spread of COVID-19. Engebretsen et al. (2020) developed a probabilistic spatiotemporal model to simulate the spatial spread of influenza using mobile phone-derived mobility data in Bangladesh. Furthermore, Tao Hu et al. (2021) conducted a comprehensive review on the use of human mobility data in COVID-19-related research, summarizing data sources, characteristics, applications, and future challenges.

1.2 Challenges in Using Existing Mobility Data on Minors for Infectious Disease Control

As demonstrated by the aforementioned studies, understanding the dynamic nature of human mobility is essential for formulating effective policies aimed at controlling infectious diseases. Moreover, as previously noted, it is crucial to analyse and comprehend the mobility patterns of minors, who may act as potential hotspots for infection transmission.

In Japan, the mobility patterns of minors have been examined using data sources such as the Person Trip Survey (Ministry of Land, Infrastructure, Transport and Tourism, n.d.) and Pseudo People Flow data (Center for Spatial Information Science, 2023). The Person Trip Survey focuses on urban travel behaviour and collects detailed information on both household and individual attributes, as well as daily movement patterns. This allows researchers to determine who is traveling, for what purpose, from where to where, at what time of day, and by what mode of transportation. Since its first large-scale implementation in the Hiroshima metropolitan area in 1967, the survey has been conducted in major metropolitan regions across Japan and has been used to assess current urban transportation conditions, forecast future travel demand, and support the development of urban transportation master plans. However, the survey is limited to selected urban areas and is typically conducted only once every ten years in major metropolitan regions. Furthermore, survey items are not standardized across different regions, reducing the comparability of data. Because the survey relies on direct distribution of questionnaires, recent challenges include declining response rates due to increased public concern over privacy and fraud, resulting in higher operational costs.

Pseudo people flow data, on the other hand, reconstructs typical 24-hour weekday travel patterns across Japan by integrating open survey data with commercially available, low-cost datasets. However, since pseudo people flow data is generated through agent-based simulation, it has inherent limitations in terms of accuracy. Furthermore, because it represents only a single weekday, it fails to capture long-term behavioural trends or mobility patterns tied to specific dates or time periods, thereby limiting its usefulness for infectious disease modelling over extended time frames.

Mobility surveys that include minors have also been conducted in countries outside Japan (Department for Transport, 2025; Statistics Netherlands, n.d.; U.S. Department of Transportation, n.d.). However, these studies are typically carried out infrequently, and their high implementation costs make it difficult to respond swiftly to societal changes or public health emergencies. Moreover, as these surveys are typically self-reported, they place a considerable burden on respondents, resulting in lower participation rates and limitations in data accuracy.

1.3 Purpose of This Study

In light of the aforementioned challenges, human mobility data derived from mobile device location histories is expected to contribute to a more precise understanding of minors' movement patterns. However, according to the Children and Families Agency's FY2023 Youth Internet Usage Environment Survey, the smartphone ownership rate among children under the age of nine in Japan is only 15.9%, indicating a very low level of personal device possession (Children and Families Agency, 2024). Similarly, in many other countries, children typically do not own personal devices until around the age of ten (Perowne and Gutman, 2023; Marketing Charts, 2022). Consequently, it remains difficult to capture minors' mobility solely through mobile-based location data.

This limitation in capturing minors' mobility through mobile-based location data is not limited to infectious disease control; it also presents challenges in other areas of public policy, such as child-rearing support and urban planning. In fact, Japan's Ministry of Land, Infrastructure, Transport and Tourism (2025) highlighted this issue in its 2025 Case Studies on the Utilization of Human Mobility Data, noting that the difficulty of obtaining mobility data for minors has posed constraints on policy development in municipalities such as Imizu City in Toyama Prefecture and Ube City in Yamaguchi Prefecture.

Despite these data limitations, it is important not to overlook the predictable nature of minors' daily activities. Their activity ranges are typically limited—primarily concentrated around their residences, schools, or childcare facilities. Due to the highly structured and repetitive nature of their daily routines, it remains feasible to estimate their mobility patterns with a reasonable degree of accuracy.

Against this backdrop, the objective of the present study is to develop a method for estimating the behavioural patterns and spatial activity ranges of minors residing in individual households across Japan by integrating various existing statistical datasets.

As a result, it becomes possible to construct mobility data for minors. Minors serve as vital connectors between households, schools, and local communities. Therefore, the mobility data developed in this study is expected to serve not only as a valuable resource for infectious disease control but also as foundational information applicable across various domains of public policy.

For instance, understanding how minors use public facilities can contribute to the planning, enhancement, and safety of urban infrastructure. Furthermore, by analysing commuting routes and after-school movement patterns, this dataset can inform concrete policy measures aimed at improving school route safety and traffic environments. Ultimately, such insights are expected to support the realization of a more sustainable society that effectively responds to the needs of minors.

As a first step, this study estimates the residential locations of minors, the educational institutions they attend, their daily schedules, and their commuting routes. Through these estimations, this study establishes a method for capturing the mobility patterns of individual minors on a nationwide scale. By enabling the analysis of detailed behavioural patterns while ensuring the protection of personal information, this approach represents a novel and distinctive contribution to the field.

2. Methodology

The flow of this study is shown in Figure 1. As Figure 1 illustrates, this study first developed a building-level population estimation dataset, referred to as the Micro Population Census (MPC). The original five-year age groupings in the MPC were disaggregated into single-year intervals, enabling the extraction of individuals under the age of 18. These data were then reorganized by school grade level to produce grade-specific population estimates for minors. Next, each minor in the grade-specific MPC was linked

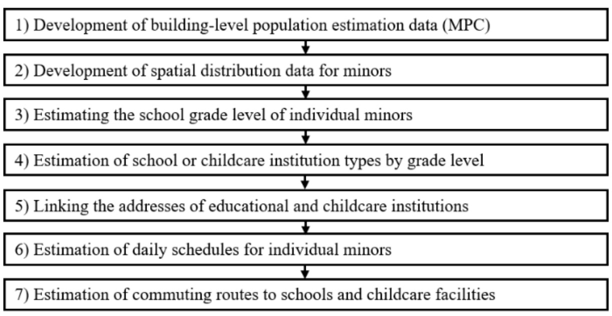

Figure 1. Flow of this study

to the address of the school or childcare facility they were presumed to attend. Then, by estimating each individual's daily schedule and commuting route, we constructed a dataset capable of representing the routine behavioural patterns and spatial activity ranges of individual minors.

2.1 Development of Building-Level Population Estimation Data (MPC)

Various demographic statistics, including the national census, are widely used as foundational data by government ministries and agencies and play a critical role in academic research across disciplines. In the private sector, such data are frequently utilized to inform business strategy, underscoring the indispensable role of demographic information in society.

However, existing demographic statistics have a notable limitation: within survey units that span both densely and sparsely populated areas, population distribution is often homogenized. This poses challenges for accurately identifying fine-grained population estimates, which are essential for designing effective infectious disease control measures.

To address this issue, the present study disaggregates census data onto residential maps that identify individual building locations. This approach enables the construction of a building-level, non-aggregated demographic dataset. In this study, we refer to this dataset as the Micro Population Census (MPC).

2.1.1 MPC Construction Method

This study developed a building-level demographic dataset incorporating attributes such as household distribution, household size, family composition, and the estimated age and gender of household members, based on the methodology proposed by Akiyama et al. (2013). Figure 2 illustrates the flow of the MPC construction method. Specifically, we first extracted buildings likely to be residential using digital residential maps, generating a "candidate household location dataset" in accordance with the criteria outlined in Table 8 of the census small-area aggregation.

Subsequently, Number of persons per household (hereinafter referred to as "Household sizes") were assigned to these candidate buildings based on the distribution presented in Table 5 of the census, ranging from single-person to seven-person households. Feasible family compositions were then allocated according to the assigned household sizes. Table 1 summarizes the various family structure types and their corresponding household sizes. The family structure types comprehensively

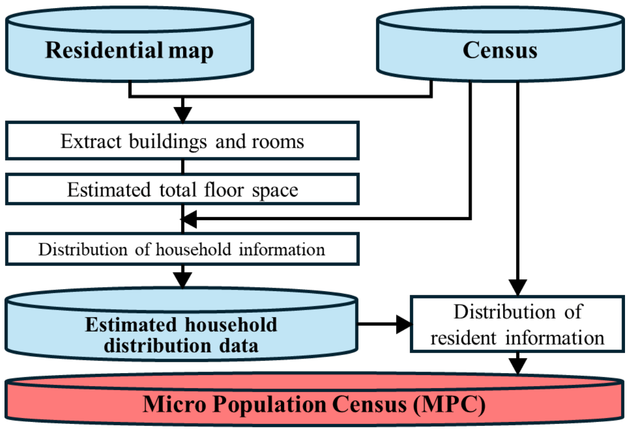

Figure 2. Flow of MPC Construction Method

capture and classify household compositions in Japan. Finally, the age and gender of household heads and members were estimated based on the assigned family composition.

Compared to the method proposed by Akiyama et al. (2013), the approach in this study introduces several modifications aimed at improving estimation accuracy. Specifically, multiple statistical datasets with shared aggregation units were utilized. Household information was probabilistically assigned to satisfy one of these datasets, allowing for the generation of detailed demographic estimates at the building level. The data were then re-aggregated according to the units of the remaining statistical datasets. Where discrepancies emerged, a substitution process was applied between conflicting units to minimize overall errors across the multiple datasets.

2.1.2 Results of MPC Construction

Figure 3 illustrates the estimated distribution of household members at the building level in the central area of Maebashi City, Gunma Prefecture. The color gradient visualizes the estimated number of residents per building, enabling the spatial variation in residential density and the clustering of housing areas to be clearly observed. This figure demonstrates that population estimates are provided for each individual building, allowing for a detailed understanding of population distribution at a microscale. Similarly, family composition, along with the age and gender of household heads and members, was estimated at the building level nationwide.

Furthermore, as the MPC includes precise location information, it can be seamlessly integrated with other existing statistical datasets. Owing to this characteristic, the dataset serves as a highly useful foundation for this study, which aims to estimate the mobility of minors by integrating various statistical sources.

2.1.3 Reliability Assessment of the MPC

Table 2 presents the results of regression analyses conducted between the MPC—re-aggregated by household size, household composition, and age–gender population categories at the small-area level—and the corresponding original data from the 2015 national census. High levels of agreement were confirmed across all prefectures, indicating that the MPC demonstrates a high degree of reliability.

2.2 Development of Spatial Distribution Data for Minors

Based on the MPC constructed in Section 2.1, we extracted minors and developed a dataset that enables the spatial distribution of minors to be identified at the building level. Since the MPC is originally aggregated in five-year age groups, we disaggregated these into single-year intervals to isolate individuals under the age of 18. Using nationwide age- and gender-specific population estimates published by the Statistics Bureau of Japan, we calculated the age composition ratios within each five-year group. These ratios were then used to proportionally assign a specific age to each individual in the MPC, enabling age estimation on a one-year basis. Individuals under the age of 18 were then extracted to develop spatial data representing the distribution of minors.

Figure 4 illustrates the estimated spatial distribution of minors at the building level for a selected area of Okayama City, Okayama Prefecture. Each plot represents the age group of minors residing in individual buildings, allowing for the visual identification of residential density and spatial variation in age composition across urban areas.

Family structure types	Household sizes
A. Household consisting only of relatives	
1. Nuclear family household	
(1) Married couple only	2
(2) Married couple with children	3~
(3) Single-father with children	2~
(4) Single-mother with children	2~
2. Non-nuclear family households	
(5) Married couple with both parents	4
(6) Married couple with one parent	3
(7) Married couple with children and both parents	5~
(8) Married couple with children and one parent	4~
(9) Married couple with other relatives (excluding parents and children)	3~
(10) Married couple with children and other relatives (excluding parents)	4~
(11) Married couple with parents and other relatives (excluding children)	4~
(12) Married couple with children, parents, and other relatives	5~
(13) Siblings only	2~
(14) Others (not classified above)	2~
B. Households including non-relatives	2~
C. One-person households	1
Household family type: unknown	1~
Institutional households	

Table 1. Classification of family types and corresponding household sizes

Figure 3. Estimated population per building unit based on MPC (example of the city center of Maebashi City)

2.3 Estimating the School Grade Level of Individual Minors

To estimate the educational institutions (schools or childcare facilities) attended by minors, we first estimated each individual's school grade based on age. Because the MPC reflects the

Prefecture	Household size -based household count		Family structure -based household count		Age and gender -based population	
	R^2	MAE	R^2	MAE	R^2	MAE
Hokkaido	0.994	0.153	0.960	0.633	0.976	1.420
Aomori	1.000	0.004	0.942	2.542	0.975	2.138
Iwate	1.000	0.009	0.931	2.878	0.975	2.238
Miyagi	1.000	0.010	0.931	3.124	0.967	2.870
Akita	1.000	0.005	0.923	2.355	0.973	1.782
Yamagata	1.000	0.005	0.901	2.944	0.967	2.870
Fukushima	1.000	0.006	0.944	2.355	0.973	1.782
Ibaraki	1.000	0.003	0.935	2.949	0.978	2.108
Tochigi	1.000	0.005	0.941	2.655	0.978	2.108
Gunma	1.000	0.006	0.942	4.060	0.974	3.468
Saitama	1.000	0.009	0.966	4.265	0.976	3.468
Chiba	1.000	0.006	0.964	3.927	0.977	4.627
Tokyo	1.000	0.013	0.967	5.552	0.971	2.209
Kanagawa	1.000	0.013	0.969	4.840	0.979	7.714
Niigata	1.000	0.003	0.904	2.646	0.964	2.109
Toyama	1.000	0.004	0.905	2.456	0.965	2.002
Ishikawa	1.000	0.004	0.930	2.140	0.971	2.029
Fukui	1.000	0.004	0.889	2.447	0.957	1.959
Yamanashi	1.000	0.004	0.945	3.562	0.981	2.922
Nagano	1.000	0.011	0.932	4.107	0.980	2.973
Gifu	1.000	0.003	0.936	1.926	0.977	1.727
Shizuoka	1.000	0.006	0.953	5.148	0.980	3.638
Aichi	1.000	0.006	0.953	2.404	0.978	2.712
Mie	1.000	0.007	0.941	2.770	0.973	2.862
Shiga	1.000	0.005	0.931	2.848	0.967	2.897
Kyoto	1.000	0.006	0.972	1.164	0.976	1.715
Osaka	1.000	0.009	0.976	2.773	0.967	4.484
Hyogo	1.000	0.004	0.969	2.065	0.976	2.838
Nara	1.000	0.005	0.969	1.941	0.977	2.322
Wakayama	1.000	0.006	0.971	1.852	0.982	2.071
Tottori	1.000	0.003	0.944	1.918	0.975	1.653
Shimane	1.000	0.006	0.926	4.121	0.973	2.916
Okayama	1.000	0.008	0.961	1.963	0.982	1.990
Hiroshima	1.000	0.008	0.949	2.600	0.974	3.310
Yamaguchi	1.000	0.004	0.953	1.465	0.980	1.985
Tokushima	1.000	0.004	0.949	2.728	0.984	2.250
Kagawa	1.000	0.005	0.953	5.217	0.984	4.212
Ehime	1.000	0.007	0.967	1.982	0.983	2.249
Kochi	1.000	0.007	0.976	1.402	0.985	1.722
Fukuoka	1.000	0.004	0.964	2.574	0.973	3.385
Saga	1.000	0.005	0.920	2.704	0.961	2.471
Nagasaki	1.000	0.008	0.960	2.012	0.973	2.235
Kumamoto	1.000	0.007	0.942	2.680	0.977	3.222
Oita	1.000	0.007	0.957	2.424	0.980	2.662
Miyazaki	1.000	0.009	0.972	1.790	0.983	2.253
Kagoshima	1.000	0.009	0.978	2.389	0.982	3.862
Okinawa	1.000	0.010	0.972	4.319	0.961	6.129

Note: All results were statistically significant at the 5% level.

Table 2. Comparison of the national census and re-aggregated MPC

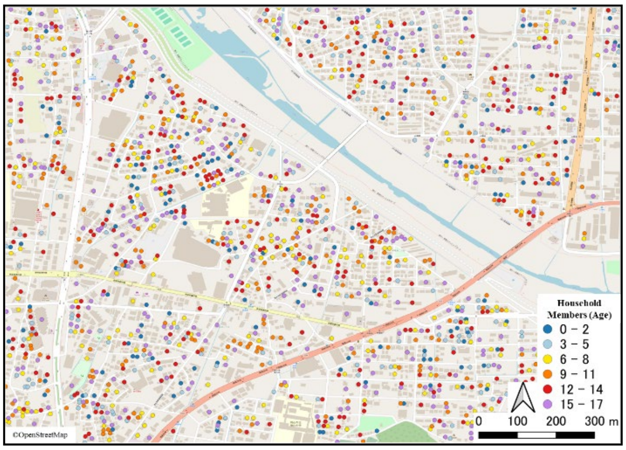

Figure 4. Estimated spatial distribution of minors at the building level (example of part of Okayama City)

population as of the census reference date—October 1—minors within the same grade level typically span two adjacent ages, with approximately half of the students differing in age by one year. To account for this, we randomly divided minors within each age group in the MPC into two equal subsets: one half was assigned to the grade corresponding to their age, while the other was assigned to the grade one year below. This probabilistic approach allowed for the estimation of each minor's likely school grade.

2.4 Estimation of School or Childcare Institution Types by Grade Level

Using the school grade classifications derived in Section 2.3, we then estimated the type of educational or childcare institution each minor was likely to attend. The categories considered in this study include not enrolled in any institution (pre-preschool), nursery school, kindergarten, certified child centre (kodomo-en), elementary school, junior high school, and high school.

For preschool-aged children not yet enrolled in elementary school, two data sources were used to estimate age-specific enrolment distributions: (1) preschool enrolment rates by age (for children not yet enrolled, kodomo-en, kindergartens, and nursery schools) from the Children and Families Agency, and (2) prefecture-level enrolment counts for kindergartens and kodomo-en institutions from the School Basic Survey. Minors were then proportionally allocated to each institution type by age. For those at the elementary, junior high, and high school levels, all minors were directly classified into their respective institution types based on grade.

2.5 Linking the Addresses of Educational and Childcare Institutions

Based on the institutional classifications assigned in Section 2.4, we linked each minor to the most likely school or childcare facility by address. First, using prefecture-level public and private student enrolment data from the School Basic Survey, we calculated the proportion of minors attending public versus private institutions and allocated individuals accordingly.

Next, we used Telepoint Data, a geocoded telephone directory provided by Zenrin Co., Ltd., to identify facilities whose names included the term "nursery school" (hoikuen). For minors categorized as nursery school attendees, the address of the nearest such facility was linked. Similarly, for those attending kodomo-en, kindergartens, high schools, private elementary schools, or private junior high schools, school address data from the National Land Numerical Information (NLNI) database were used to assign the nearest institution within each category.

For minors attending public elementary and junior high schools, school district boundary data from the NLNI were used to determine the appropriate school based on residential location. However, due to gaps in the school district dataset, minors who could not be matched to a school through district data were instead assigned the nearest public school using the same proximity-based method applied to other institution types.

2.6 Estimation of Daily Schedules for Individual Minors

We then estimated the daily schedules of individual minors. For each educational or childcare institution type, reasonable arrival and departure times were defined based on standard attendance patterns. Subsequently, the straight-line distance between each minor's residence and their assigned facility was calculated.

Assuming that longer commutes involve faster transportation modes, we assigned travel speeds as follows: walking (4 km/h) for distances up to 1 km; bicycle (10 km/h) for 1–5 km; and automobile (30 km/h) for distances exceeding 5 km. Morning departure times from home were calculated by subtracting the estimated travel time from the arrival time, while return times were calculated by adding the travel time to the departure time from the facility.

2.7 Estimation of Commuting Routes to Schools and Childcare Facilities

Finally, we estimated the commuting routes for each minor. Specifically, we utilized origin-destination (OD) coordinate pairs representing home and educational facility locations, derived in Section 2.5, in conjunction with the inferred travel modes identified in Section 2.6. Route estimation was conducted using Google Maps Routes API (Compute Routes).

This API supports multiple travel modes—including walking, bicycling, and driving—and calculates optimal routes according to the characteristics unique to each mode. For walking and bicycling, the API generates paths along pedestrian- or bicycle-appropriate roads and pathways, as defined within the underlying road network data. For driving, it accounts for traffic conditions, network topology, one-way restrictions, and other regulatory constraints to compute time-efficient routes.

Notably, in the case of bicycling, the API may fail to return a complete route to the destination when dedicated bicycle lanes are unavailable, resulting in truncated paths. To ensure consistency and completeness in route generation, the walking mode was uniformly applied to both walking and bicycling travel modes.

This approach enabled the estimation of realistic travel distances, durations, and route geometries for each minor, aligned with their respective inferred travel modes.

3. Results

Using the proposed method, we were able to estimate, on a nationwide scale, the school or childcare facility attended by each minor residing in an individual household, along with their estimated departure and return times and commuting routes.

Figure 5 illustrates the distribution of elementary school students' residences and their corresponding public and private schools in the Yokohama City area of Kanagawa Prefecture. In the figure, the residential locations of students attending public and private elementary schools are visualized as origin–destination (OD) pairs, connecting their place of residence (origin) with their assigned school (destination). Clear differences can be observed between public and private school commuting patterns in terms of both geographic range and density. Specifically, students commuting to private schools are distributed across a wider area, whereas students attending public schools generally commute within their designated school districts. The spatial relationship between students' homes and schools also enables visual assessment of school district boundaries and actual commuting distances. This provides valuable baseline information for evaluating the spatial distribution of educational facilities and the geographic disparities in commuting burdens across regions. However, as noted in Section 2.5, school assignments were determined algorithmically based on proximity. As a result, in cases where no suitable facilities exist nearby, students were assigned to distant institutions, sometimes even across municipal boundaries, which may not reflect realistic commuting behaviour.

Figure 6 visualizes the estimated departure times from home for students commuting to a specific elementary school. Each student's home and school are connected by a line representing the origin–destination (OD) pair, with the lines color-coded according to departure time intervals (e.g., 7:45–8:30). This visualization enables the identification of patterns in departure times based on commuting distance, as well as the temporal spread of student movement during the school arrival period.

Figure 7 visualizes the estimated school commuting routes of students attending an elementary school in Kashiwa City, Chiba Prefecture. The estimated routes, represented by blue lines, extend from each student's residential location to the school. The visualization reveals a radial pattern of commuting paths centered on the school, and a directional concentration of routes that reflects the spatial distribution bias of student residences.

This type of spatial visualization facilitates the identification of high-risk areas along commuting routes and zones where traffic safety interventions are most needed. As such, it provides valuable information to support decision-making by boards of education and local governments in planning and implementing student commuting assistance and safety measures.

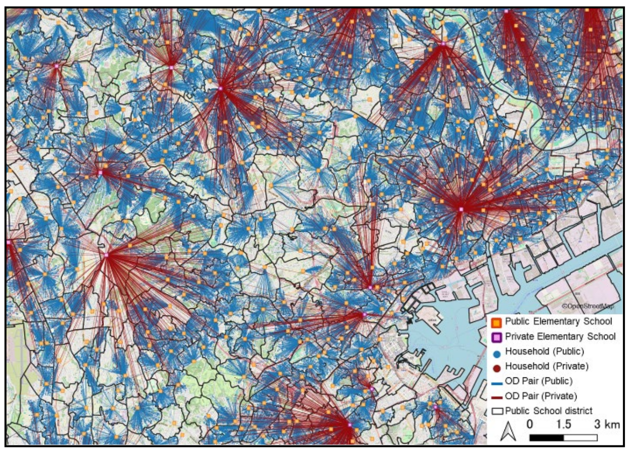

Figure 5. Distribution of elementary school students and their schools (example from the Yokohama Area)

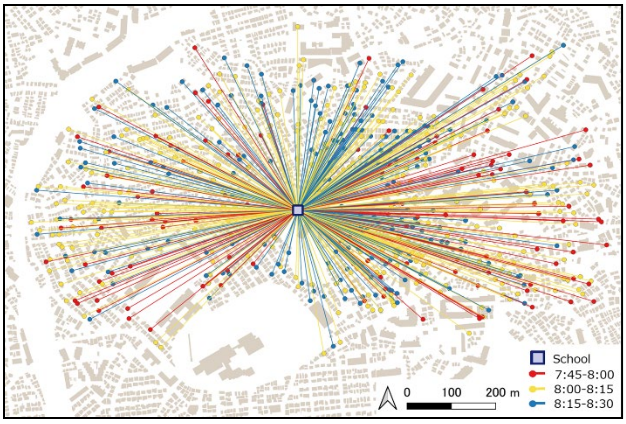

Figure 6. Estimated departure times for students attending a specific elementary school

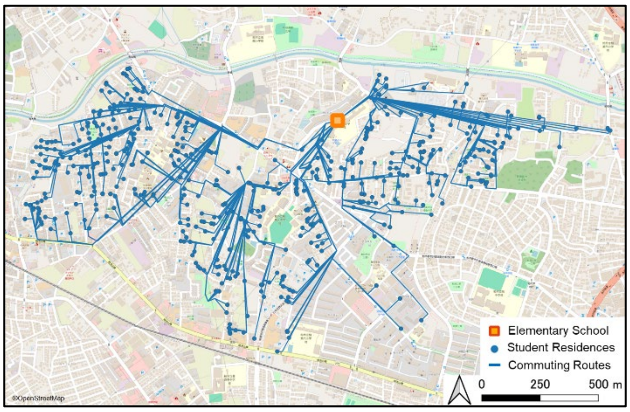

Figure 7. Estimated commuting routes of elementary school students (example from a school in Kashiwa City)

It should be noted, however, that as described in Section 2.7, the analysis assumes that commuting modes are limited to walking, bicycling, or travel by car. This simplification may introduce discrepancies between the estimated routes and actual commuting behaviours.

Based on these results, the method developed in this study enables the nationwide estimation of each minor's place of residence, school or childcare facility, daily schedule, and commuting route. These data collectively offer a fundamental framework for representing the routine behavioural patterns and spatial activity ranges of minors across Japan.

4. Conclusion

This study developed a method for estimating the school or childcare facility attended by each minor, along with their daily schedules and commuting routes, in order to capture the spatial distribution and routine behavioural patterns of minors at the building level on a nationwide scale.

To enhance the accuracy of the model, future work will incorporate estimated commuting times into the school and childcare facility allocation process. If the estimated travel time is deemed unrealistic, students will be reallocated to alternative institutions attended by peers of the same age group, thereby ensuring more plausible assignments.

Furthermore, we plan to incorporate public transportation as a potential travel mode by utilizing data sources such as the Navitime API and GTFS. We also aim to account for multimodal commuting patterns, including combinations of walking and train travel.

In addition, we seek to improve the classification accuracy of school destination types by utilizing small-area income statistics and considering the existence of single-gender schools. We also intend to analyze non-school-commute mobility patterns—including visits to after-school care programs, extracurricular activities, parks, shopping areas, and other destinations—to further refine the estimation of minors' behavioral patterns and activity spaces.

Looking ahead, we aim to conduct more rigorous reliability validation. Although acquiring the necessary data presents certain challenges, we plan to evaluate the population distribution generated by the MPC and assess its accuracy at finer spatial scales. To further examine the mobility patterns of minors, we also intend to utilize GPS data from child monitoring devices, as well as school commute route data obtained from the National Police Agency and individual educational institutions.

Building upon this foundation, we will integrate the dataset developed in this study with human mobility data derived from mobile device location histories. This integration will enable the construction of a comprehensive, real-time dataset that captures the mobility of all age groups. Consequently, such a dataset could make a significant contribution to the design and implementation of more effective infectious disease control strategies.

For example, we envision the development of an application capable of simulating scenarios in which, following the confirmation of an infection at a specific school, a certain proportion of children transmit the infection to their households, and a portion of their guardians subsequently spread it in their workplaces. This system would enable the identification of locations at elevated risk of secondary transmission following an initial outbreak, thereby facilitating proactive and timely public health interventions—contrasting with the reactive measures that have often been employed in the past. Furthermore, we plan to validate the model by using empirical data recorded during actual infectious disease outbreaks.

Moreover, the human mobility dataset for minors developed in this study enables microscale identification of spatial distribution and commuting behaviours on a nationwide level. This capability holds implications not only for infectious disease control but also for a wide range of social infrastructure applications. Specifically, the dataset can support the optimal placement of childcare facilities, schools, cram schools, and commercial establishments; improve public transportation route planning; and enhance strategies for allocating disaster preparedness supplies.

From a traffic safety perspective, the dataset can also support the development of targeted measures such as the time- and location-specific assignment of traffic control personnel based on mobility patterns. Taken together, these potential applications contribute to the advancement of evidence-based policy making (EBPM) and underscore the value of incorporating minors' mobility data as a novel foundation for decision-making in urban planning and public service design.

Acknowledgements

This study is the result of joint research between LocationMind, Inc. and CSIS, the University of Tokyo (No. 990). The development of the building-based population data (Micro Population Census: MPC) used in this study was greatly supported by Naoya Takeda, who completed his master's degree at Tokyo City University in March 2023. Additionally, we extend our sincere gratitude to Ryota Yamanotera, who completed his master's degree at Tokyo City University in March 2024, for his invaluable support.

References

Akiyama, Y., Takada, H., Shibasaki, R., 2013: Development of micro population census through disaggregation of national population census. *CUPUM2013 conference papers*, 1–33.

Center for Spatial Information Science, University of Tokyo, 2023: Nationwide Pseudo People Flow Data Specification Ver.2.0. https://pflowwp.sekilab.global/wp-content/uploads/Pseudo-PFLOW-Specification-ver2.0.pdf (7 April 2025).

Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU), 2023: COVID-19 Dashboard. https://gisanddata.maps.arcgis.com/apps/dashboards/bda7594740fd40299423467b48e9ecf6 (7 April 2025).

Children and Families Agency, 2024: Survey on Internet Use Environment of Youths in FY2023. https://www.cfa.go.jp/assets/contents/node/basic_page/field_ref_resources/9a55b57d-cd9d-4cf6-8ed4-3da8efa12d63/fc117374/20240226_policies_youth-kankyou_internet_research_results-etc_09.pdf (7 April 2025).

Department for Transport, 2025: National Travel Survey. https://www.gov.uk/government/collections/national-travel-survey-statistics (7 April 2025).

Doi, K., Onishi, M., 2023: Analysis of People's Flow using Large-Scale GPS Data. *IPSJ 85th National Convention*, 67–68.

Engebretsen, S., Engø-Monsen, K., Aleem, M.A., Gurley, E.S., Frigessi, A., Freiesleben de Blasio, B., 2020: Time-aggregated mobile phone mobility data are sufficient for modelling influenza spread: the case of Bangladesh. *Journal of the Royal Society Interface*.

Hu, T., Wang, S., She, B., Zhang, M., Huang, X., Cui, Y., Khuri, J., Hu, Y., Fu, X., Wang, X., Wang, P., Zhu, X., Bao, S., Guan, W., Li, Z., 2021: Human mobility data in the COVID-19 pandemic: characteristics, applications, and challenges.

Ilin, C., Annan-Phan, S., Tai, X.H., Mehra, S., Hsiang, S., Blumenstock, J.E., 2021: Public mobility data enables COVID-19 forecasting and management at local and global scales. *Scientific Reports*, 11, Article number: 13531.

Kawakami, Y., Nojiri, S., Nakamoto, D., Irie, Y., Miyazawa, S., Kuroki, M., Nishizaki, Y., 2023: A novel indicator for the spread of new coronavirus disease 2019 and its association with human mobility in Japan. *Scientific Reports*, 13, 1–12.

Marketing Charts, 2022: At What Age Do Kids Start Getting Smartphones? https://www.marketingcharts.com/demographics-and-audiences/teens-and-younger-225502 (7 April 2025).

Ministry of Health, Labour and Welfare (MHLW), 2021: Infection Control Guidelines in Nursery Schools (Revised 2018 Edition). https://www.mhlw.go.jp/content/001005138.pdf (7 April 2025).

Ministry of Health, Labour and Welfare (MHLW), n.d.: Basic Knowledge on Infection Control | Part 1. https://www.mhlw.go.jp/content/000501120.pdf (7 April 2025).

Ministry of Land, Infrastructure, Transport and Tourism (MLIT), 2025: Case Studies on the Utilization of Human Mobility Data. https://www.mlit.go.jp/tochi_fudousan_kensetsugyo/chirikukannjoho/content/001881566.pdf (21 April 2025).

Ministry of Land, Infrastructure, Transport and Tourism (MLIT), n.d.: Urban Transport Surveys and Person Trip Surveys. https://www.mlit.go.jp/toshi/tosiko/toshi_tosiko_tk_000031.html (7 April 2025).

Mishima, Y., Habault, G., Wada, S., 2022: A study of human mobility data structure for COVID-19 cases prediction. *Proceedings of the 36th Annual Conference of the Japanese Society for Artificial Intelligence*, 1–4.

Onoda, K., Asano, R., Suzuki, T., 2020: Possibilities of Mobile Big-Data in the Age of COVID-19. *The 11th Transdisciplinary Federation of Science and Technology Conference*, 1–4.

Perowne, R., Morrison Gutman, L., 2023: Parents' perspectives on smartphone acquisition amongst 9- to 12-year-old children in the UK – a behaviour change approach. Published online: 03 May 2023.

Statistics Netherlands, n.d.: Dutch National Travel Survey. https://www.cbs.nl/en-gb/our-services/methods/surveys/brief-survey-description/dutch-national-travel-survey (7 April 2025).

U.S. Department of Transportation, n.d.: National Household Travel Survey. https://nhts2024.com (7 April 2025).