The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume L-4/W1-2026
Free and Open Source Software for Geospatial (FOSS4G) 2026 – Academic Track “Bridging Geospatial Technology and Humanity”, 30 August–5 September 2026, Hiroshima, Japan

An Open Geospatial Framework for Proximity-Based Community Planning: Integrating Mobility-Based Community Detection, Multimodal Accessibility and Facility Allocation Optimisation

Junyoung Choi[1]*, Sunyong Eom[2], Jiho Yeo[3], Yuri. Lee[1]

[1] The Seoul Institute, 57 Nambusunhwan-ro 340-gil, Seocho-gu, Seoul 06756, Republic of Korea, (junyoung.choi, yuri.lee)@si.re.kr
[2] Hanyang University, Graduate School of Urban Studies, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Republic of Korea, sunyong.eom@hanyang.ac.kr
[3] Gachon University, Department of Smart City Engineering, 1342 Seongnam-daero, Sujeong-gu, Seongnam 13120, Republic of Korea, jiho.yeo@gachon.ac.kr

Keywords: Proximity-based planning, FOSS4G, Mobility-based community detection, Multimodal accessibility, Facility allocation, Linear programming, Busan

Abstract

Proximity based community planning organises essential services so residents can reach them within short travel distances, but its implementation requires combining large scale mobility data, multimodal accessibility modelling and spatial optimisation, components typically fragmented across different tools. We propose an open geospatial framework that integrates these three strands into one reproducible FOSS workflow. The framework has three components. The first performs mobility based community detection on telecom trip networks. The second computes multimodal accessibility on a 250 m grid, combining OSRM road routing with r5py walking and public transport routing over OpenStreetMap and GTFS. The third allocates facilities by linear programming whose objective jointly considers equity and economic efficiency. We demonstrate the framework on Busan Metropolitan City, covering 16 districts plus Gijang County, 770 km², 12,474 grid cells and 4,119 facilities across 18 categories. Community detection returns 50 stable living area communities derived from accessibility and 50 derived from mobility, with generalised modularity 0.918 and 0.934, mean best IoU 0.502, ARI 0.547 and NMI 0.814. Applied to a sample sub region of five inner city districts with 820,344 residents, the LP places 15 new facility units across 9 candidate 250 m grid cells and lifts walking only coverage of the five most deficient service types by 12 to 20 percentage points each.

1. Introduction

Cities are shifting from car-oriented, broad-scale planning to ‘daily living areas’ organised around what residents can do on foot within a short time budget (Moreno et al., 2021; Allam et al., 2022). This is not a stylistic trend but a sustainability strategy: walking and cycling reduce carbon emissions, improve health, revitalise neighbourhood commerce and strengthen social ties (Choi, 2025). In Korea, the 2024 amendment of the National Land Planning and Use Act has elevated the Living-Area Plan (Saenghwalgwon Gyehoek) to statutory status; municipalities are newly obliged to delineate walkable living-area boundaries and to assess essential-service supply within them in a quantitative and reproducible manner.

Traditional living-area plans in Korea were drawn along administrative boundaries (a dong–eup–myeon scale for everyday living areas, and an autonomous-district scale for regional living areas). The approach was easy to administer but routinely misaligned with how people actually move: terrain, rivers, transit corridors and commuting patterns rarely respect district lines, so a single administrative ‘living area’ often bundles disconnected neighbourhoods or splits tightly interacting ones (Choi, 2025). Operationalising a more truthful, proximity-based plan therefore requires an analytical pipeline that integrates large-scale mobility data, multimodal accessibility modelling and spatial optimisation. In many cities these three components remain fragmented across different tools and datasets, which obstructs reproducible and scalable workflows.

Mature components of the FOSS4G ecosystem now make such an integrated pipeline practical: open routing engines such as OSRM (Luxen and Vetter, 2011) and the R5 stack accessed through r5py (Tenkanen et al., 2022; Pereira et al., 2021), open graph and community-detection libraries (NetworkX, python-louvain), and free MILP solvers such as COIN-OR CBC. Python libraries for spatial data handling, network analysis and walkable-network accessibility have lowered the technical barrier to complex urban analysis (Choi, 2025), and this openness democratises the planning process: small municipalities, academia and civil society now have the same analytical leverage as large institutions.

This paper proposes a three-component open geospatial framework for proximity-based community planning. First, mobility-based community detection identifies functional urban communities from telecom origin–destination (OD) flows. Second, multimodal accessibility analysis combines pedestrian routing via OSRM with walking-plus-transit routing via r5py to compute supply-and-demand-aware accessibility on a 250 m grid. Third, a linear-programming (LP) facility-allocation module evaluates where to consolidate, place or convert essential-service facilities, balancing efficiency and equity (Choi, 2025). The framework is implemented as the Living-Zone Dashboard, demonstrated on Busan Metropolitan City - the case study selected for the 2025–2026 pilot of the Korean R&D project ‘Big-Data-Based AI Urban-Planning Technology’ (Seoul Institute, 2024). The dashboard runs end-to-end on 12,474 KOSTAT 250 m grid cells covering the 16 districts and Gijang County of Busan (3.15 million residents) with a ledger of 4,119 facilities across 18 essential-service categories. Section 2 reviews related FOSS4G-relevant work; Section 3 develops the three-component methodology; Section 4 describes the case study, the ethics-and-privacy protocol and the software stack; Section 5 reports quantitative results on the full Busan grid and on a sample inner-city sub-region (Nam-gu, Busanjin-gu, Dong-gu, Jung-gu and Yeongdo-gu; 820,344 residents); Section 6 discusses implications and limitations; Section 7 concludes.

2. Related Work

2.1 Proximity, walkability and international living-area cases

Proximity-based planning frames urban quality of life around what residents can do within short walking, cycling or transit distances of home (Moreno et al., 2021; Allam et al., 2022). Walkability itself is a multidimensional capacity of urban form - best captured by what Dovey and Pafka (2019) call ‘urban DMA’: Density (more people and places within walking distance), functional Mix (live-work-visit synergy), and Access (the connective fabric that lets pedestrians move between them). Major cities are translating this idea into formal policy: Paris with the 15-minute city, Melbourne with 20-minute neighbourhoods, Bogotá with Barrios Vitales, Barcelona with Superblocks, and Japan with the Community-based Integrated Care System (Choi, 2025). Two implementation lessons recur across these cases: (i) data on actual movement, not administrative geography, has to drive the boundary work, and (ii) ageing-society contexts require explicit modelling of access to health, care and welfare facilities by age cohort (Choi, 2025).

2.2 Mobility-based community detection

Community detection on a spatial-interaction graph delineates functional urban units from observed flows rather than administrative fiat. In urban-scale FOSS practice, two algorithms have become standard (Choi, 2025): the Louvain algorithm, which merges nodes hierarchically to maximise modularity on undirected weighted networks, and the Infomap algorithm, which compresses random-walk trajectories on directed networks (e.g. commuter flows). The two are complementary - Louvain for undirected proximity and interaction, Infomap for directional flows. We use the python-louvain implementation over NetworkX, augmented by Eom’s (2025) α-sweep procedure with a multi-criterion selection rule (generalised modularity Q_γ, mean normalised mutual information, and a consensus score) to defuse the resolution-limit problem. Yeo et al. (2025) construct the underlying interaction graph from telecom (LGU+) trip-chain data and show that resulting mobility-derived communities deviate meaningfully from administrative boundaries; Lee et al. (2025) report similar findings comparing accessibility-based and flow-based delineations.

2.3 Open-source multimodal accessibility tools

Multimodal accessibility computation has matured rapidly in the FOSS4G ecosystem. OSRM (Luxen and Vetter, 2011) provides fast road and walking routing on OpenStreetMap. The R5 routing engine, accessible through r5py (Tenkanen et al., 2022) and its R counterpart r5r (Pereira et al., 2021), integrates the pedestrian network from OSM with public-transport schedules from GTFS feeds. The two engines are typically used in isolation; reconciling them and embedding their outputs directly into a downstream optimisation pipeline remains an open practice gap that we address in this paper (Choi, 2025).

Once routing has produced a travel-time matrix, the per-cell accessibility indicator we adopt is the two-step floating catchment area (2SFCA) family of measures, in the three-step extension of Wan, Zou and Sternberg (2012). 2SFCA is preferred over simple shortest-distance or gravity-only indicators because it embeds both facility capacity and demand population within the walking threshold - a facility that is physically nearby but over-subscribed yields a lower effective accessibility than one that is slightly further but lightly used (Choi, 2025). The framework remains indicator-agnostic: gravity or isochrone-area measures can be substituted at this stage.

2.4 Facility-location optimisation with FOSS solvers

Three families of methods are commonly available for spatial facility-allocation problems (Choi, 2025): mathematical (linear and mixed-integer) programming, which offers proven optimality and explainable solutions; heuristic and meta-heuristic methods (GA, SA, TS), which are flexible for irregular problem structures but lack optimality guarantees; and deep-learning / reinforcement-learning methods, which scale to very large search spaces but obscure the decision logic. For public-sector planning, where transparency and auditability are essential, linear programming offers the most defensible combination of mathematical rigour and interpretability. We therefore implement Component 3 as a mixed-integer LP solved with COIN-OR CBC through the PuLP interface (Mitchell et al., 2011). The Korean policy-design version of this argument - including the contrast between an efficiency-focused consolidation model for shrinking municipalities and an equity-focused placement model for growing cities - is laid out in detail in Choi (2025), and the present paper provides the corresponding empirical FOSS4G implementation.

3. Methodology

3.1 Three-component overview

Figure 1 summarises the three-component architecture. Component 1 delineates functional urban communities from telecom OD flows. Component 2 computes multimodal accessibility on a 250 m grid using OSRM and r5py. Component 3 solves a linear-programming facility allocation model whose objective function combines equity and economic efficiency. The three components communicate through tabular and raster artefacts in open formats (GeoParquet, GeoTIFF, CSV) so that every intermediate result can be inspected, replaced or recomputed by third parties - a deliberate FOSS reproducibility property. The pipeline is orchestrated by Snakemake (Köster and Rahmann, 2012) and shipped as a Docker image so that the entire workflow can be reproduced on any city whose data follow OSM, GTFS and population-grid schema

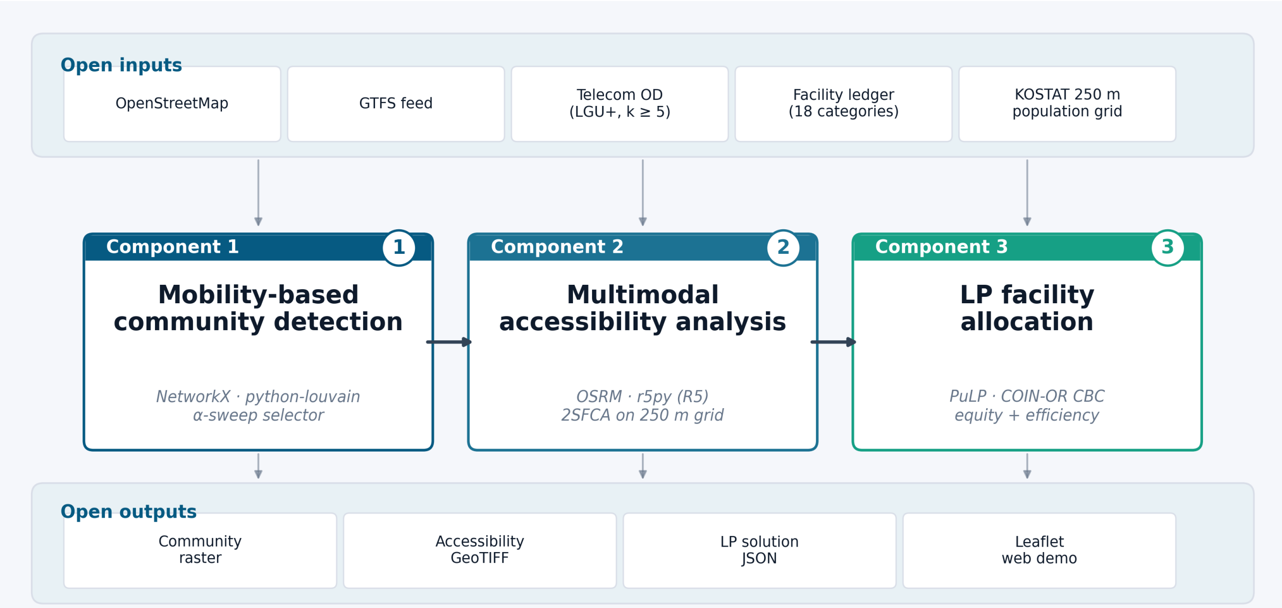

Figure 1. Three-component architecture of the open geospatial framework. All components run on FOSS; arrows indicate the canonical data flow.

3.2 Component 1 - Mobility-based community detection

Component 1 turns observed travel flows into living-area boundaries. We start with the LGU+ telecom trip-chain dataset and aggregate the trips onto the KOSTAT 250 m population grid, so that every grid cell becomes a node in a network and every edge weight is the average daily number of trips between two cells. Intuitively, two cells are 'close' in this network when many residents move between them - not when they share a postal code. We then ask the question: which clusters of cells exchange most of their daily traffic among themselves, and least with the outside? That is exactly what a community-detection algorithm answers (Choi, 2025).

We use the Louvain algorithm (Blondel et al., 2008) in python-louvain over NetworkX (Hagberg et al., 2008). In a single sentence, Louvain repeatedly merges neighbouring nodes whenever the merger raises an overall 'inside-the-cluster traffic vs. outside-the-cluster traffic' score (modularity). Because the score has a known resolution issue (it can over-merge small clusters into one giant blob), we run an α-sweep - the same procedure but with a sequence of resolution dials $\alpha \in [0.1, 2.0]$ - and pick the dial setting where two independent realisations of the algorithm agree on the same number of stable communities after removing islands smaller than 35 cells. The selection rule combines a modularity-style score Q_γ, normalised mutual information across repeated runs, and a consensus score (Eom, 2025). For the Busan-wide grid this procedure picks $\alpha = 0.85$ for the mobility graph and $\alpha = 0.80$ for the accessibility graph, both yielding 50 stable communities ($Q_\gamma = 0.934$ and 0.918, ARI = 1.0 across candidate runs - i.e. fully deterministic at the chosen resolution). The empirical α-sweep curve is shown in Figure 2.

3.3 Component 2 - Multimodal accessibility analysis

Component 2 measures how easily residents can reach each type of essential service. Two open routing engines are run in parallel on the same 250 m grid. OSRM 5.27 (Luxen and Vetter, 2011) handles pedestrian routing on the OpenStreetMap road network at 4.8 km · h^{-1}. r5py (Tenkanen et al., 2022) - the Python interface to the R5 engine, paralleling the r5r interface for R (Pereira et al., 2021) - handles walking-plus-transit routing using the OSM pedestrian network and the 2025 GTFS feed for Busan's buses, urban rail and light rail. The walking-only path is the default; the transit-augmented path is used wherever the planner wants to extend the service area into pockets that walking alone cannot cover (Choi, 2025).

The travel-time matrix is then collapsed into a per-cell accessibility score with the two-step floating catchment area (2SFCA) family of measures (Wan et al., 2012). The intuition is simple: at each facility we first compute a supply/demand ratio (how much capacity that facility has per person within its walking catchment), and then at each grid cell we sum those ratios across every facility a resident can reach. A facility that is nearby but over-subscribed contributes a small term to the sum; a facility that is slightly further but lightly used contributes a larger one. The per-cell score (Equation 1) is therefore a single number that combines proximity, capacity and demand.

$$A_i = \Sigma_j R_j \cdot f(d_{ij}) \qquad (1)$$

Here R_j is the supply-to-demand ratio computed at facility j, d_{ij} is the multimodal travel time from cell i to facility j, and f(d) is a distance-decay function - we use $f(d) = d^{-2}$ with a walking threshold of 500 m for a 15-minute living area (833 m for a 20-minute one). Per-category scores A_i^c are computed for each of the 18 service categories, and a category-weighted aggregate feeds Component 3. The implementation uses NumPy and SciPy sparse matrices and scales to the entire Busan grid without specialised hardware.

3.4 Component 3 - Linear-programming facility allocation

Component 3 is where the framework moves from diagnosis to prescription. Given the current locations of each facility type and a planner-supplied set of candidate sites (e.g. publicly owned parcels), the LP chooses which existing facilities to retain (x_j = 0 or 1), which candidate sites to open (z_j = 0 or 1), and which cell is served by which facility (y_{ij} = 0 or 1). Every decision is confined inside the same mobility-derived community from

Component 1, so the LP cannot place a service in a community whose residents would have to leave their living area to reach it. The objective function combines three terms with planner-set weights:

$$Cost(retain) + Travel(serve) - Equity(open\ new) \quad (2)$$

'Cost(retain)' penalises keeping unneeded facilities open, 'Travel(serve)' penalises long walking times between cells and the facilities that serve them, and 'Equity(open new)' rewards opening new facilities that close the gap for historically underserved cells. The same model can be turned in two opposite directions simply by changing the relative weights (Choi, 2025): an efficiency-oriented Facility Consolidation Model for shrinking sub-regions (heavy weight on the cost term, light weight on equity), and an equity-oriented Accessibility Maximisation Model for growing sub-regions (heavy weight on equity, light weight on cost). The LP also accepts standard planning constraints - every cell is served exactly once, capacity is respected, the total number of facilities is bounded, every category has a minimum count, and selected candidate sites are spatially separated (Choi, 2025).

The MILP is solved with COIN-OR CBC via the PuLP interface, in two operating modes that ship with the Living-Zone Dashboard: a target-first mode ('maximise the number of facility types that reach a planner-set coverage target, e.g. 90 % of residents within a 15-minute walk') and a capacity-gap-min mode ('minimise the residual shortfall summed across all types'). CBC returned an Optimal status for every instance run in the case study below, and the solver-abstraction layer in PuLP permits seamless substitution with other open-source MILP solvers should larger instances require it. We deliberately retain LP-based optimisation rather than heuristics or deep learning because public-sector planning requires explainable and defensible decisions (Choi, 2025).

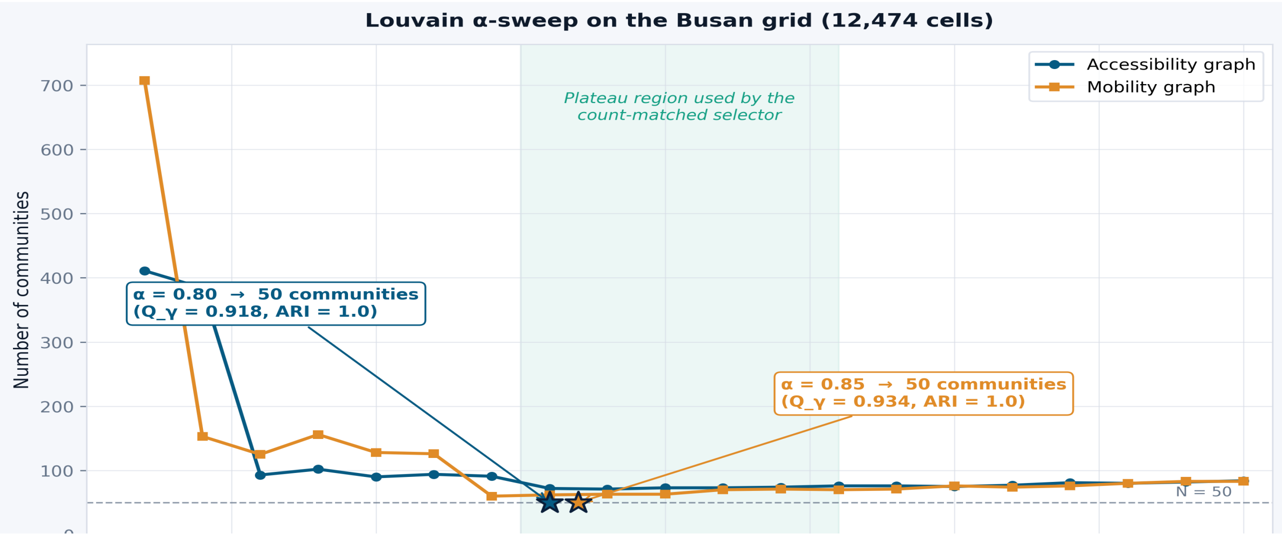

Figure 2. Empirical Louvain α-sweep on the full Busan grid (12,474 cells). The count-matched selector picks α = 0.85 (mobility) and α = 0.80 (accessibility); both converge to 50 stable communities after merging islands smaller than 35 cells.

4. Case Study

4.1 Study area and data inventory

Busan is Korea's second-largest city with 3.3 million residents on 770 km². The dashboard implementation reported here covers all 16 districts and Gijang County (3.15 million residents on the centroid-filtered 250 m grid). Five inner-city districts - Nam-gu, Busanjin-gu, Dong-gu, Jung-gu and Yeongdo-gu (820,344 residents, 11 mobility-derived communities after manual adjustment) - were selected as the sample sub-region on which Component 3 has been exercised end-to-end. These districts span from a rapidly shrinking old-town core (Jung-gu and Dong-gu) to a stable mid-sized urban district (Busanjin-gu) and ageing inner-suburban districts (Nam-gu, Yeongdo-gu). Table 1 summarises the data inputs and provenance.

Table 1. Data inventory and provenance - Living-Zone Dashboard package, Busan implementation.

Component	Dataset	Resolution / Granularity	Coverage / Volume
1	LGU+ telecom OD flows	50 m × 50 m aggregated	October 2025 monthly average
1, 2	Administrative-boundary reference	Polygon	16 districts + 1 county
2	OpenStreetMap road network	-	Busan administrative envelope
2	GTFS public-transit feed	-	Bus, urban rail, light-rail (2025)
2	T-MAP routing API	Trip-level	20 % stratified sample of OD pairs
2	KOSTAT population grid	250 m × 250 m	12,474 cells; pop 3,145,407
2, 3	Facility ledger (18 categories)	Point	4,119 facilities (Dec 2025)

4.2 Ethics, privacy and reproducibility safeguards

The framework relies on a privacy-sensitive data source - telecom-based OD flows. We adopt four explicit safeguards. First, all telecom data are obtained from LGU+ under a written research-use agreement and are delivered already aggregated to a 50 m grid with k-anonymity $k \geq 5$; cells with $k < 5$ are suppressed at source. Second, we re-aggregate to the KOSTAT 250 m grid before any downstream computation, increasing the effective k. Third, raw telecom data are processed only on a Seoul Institute internal server and are never redistributed; the public GitHub release contains only synthetic demonstration data generated by a gravity model fitted to the aggregated statistics. Fourth, the computational pipeline is logged in a Snakemake workflow with deterministic seeds, so every reported number can be traced to a specific input file and code revision - a property that aligns with both FAIR data principles and standard FOSS reproducibility practice.

4.3 Software stack and code availability

Every component runs on FOSS. Component 1 uses NetworkX 3.3 and python-louvain 0.16. Component 2 uses r5py 0.1.0 with R5 and Java 17, OSRM 5.27 and osmnx 2.x; the 2SFCA computation uses NumPy 1.26, SciPy 1.13 and GeoPandas 0.14 sparse matrices. Component 3 uses PuLP 2.7 wrapping COIN-OR CBC 2.10. Geospatial I/O is handled by GDAL 3.8 / GeoTIFF / GeoParquet. The pipeline is orchestrated with Snakemake and shipped as a Docker image; an interactive demonstration is published as a GitHub Pages site that uses Leaflet to render the produced GeoJSON layers. Code, parameter files, the Snakemake DAG and synthetic demo data will be made publicly available under an open-source licence upon publication of this paper; the repository URL will be announced at that time.

5. Results

5.1 Mobility-derived communities versus accessibility-derived communities

Figure 2 shows the empirical α-sweep on the full Busan grid. For $\alpha \leq 0.6$ both graphs produce very large numbers of communities (Louvain finds 60–700 raw communities depending on α), and small changes in α produce large changes in the partition. From $\alpha \approx 0.7$ onwards the curve flattens, and the count-matched selector identifies $\alpha = 0.85$ (mobility) and $\alpha = 0.80$ (accessibility), both returning 50 stable communities after the $m\_min = 35$ merge step. The selected partitions have $Q\_\gamma = 0.934$ (mobility) and $Q\_\gamma = 0.918$ (accessibility), with mean ARI across candidate runs of 1.0 - i.e. deterministic at the selected resolution. Figure 3 (panels a and b) maps the two 50-community partitions side by side; panel c shows the per-community best IoU between them. Mean best IoU is 0.502 (population-weighted 0.525), ARI between the two label sets is 0.547 and normalised mutual information 0.814. Of the 50 accessibility communities, 43 have an $\text{IoU} \geq 0.25$ with their best mobility match and 23 have $\text{IoU} \geq 0.50$. The two graphs agree on the broad shape of Busan’s functional structure but disagree at the boundaries - the disagreement zones are precisely the corridors along which mobility flows cross accessibility-based service catchments.

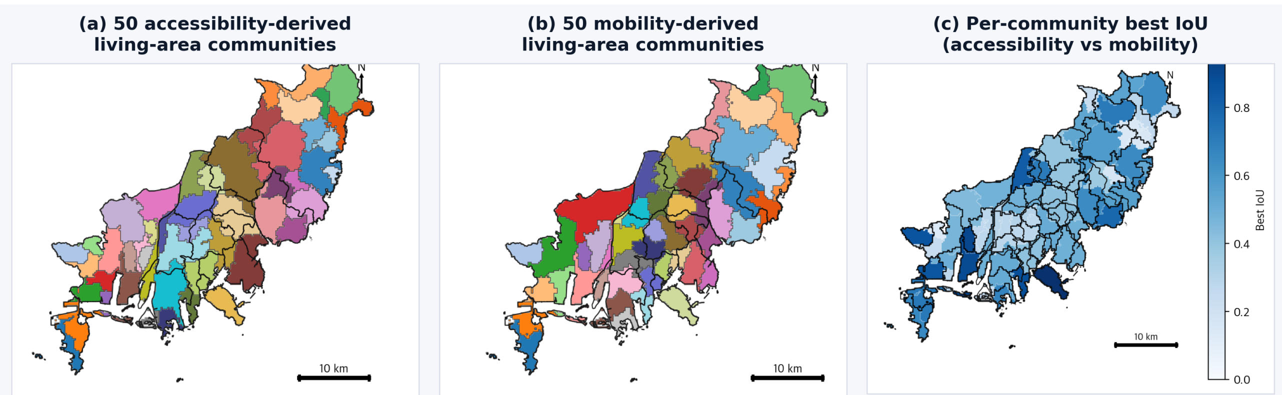

Figure 3. Empirical Busan partitions from the Living-Zone Dashboard. (a) 50 accessibility-derived communities; (b) 50 mobility-derived communities; (c) per-community best IoU between the two delineations.

5.2 Multimodal coverage of essential-service categories

Component 2 produces a per-category 2SFCA accessibility raster and, by population-weighting it, a coverage statistic. Figure 4(a) shows walking-only coverage for the 18 facility categories on the sample sub-region (820,344 residents). Schools, kindergartens and community service centres exceed the 90 % planner target; parks, exhibition halls, lifelong-learning and senior-welfare sit in the 70–80 % range; but five categories fall well below the target - sports gymnasiums (15 %), public-health centres (17 %), district offices (17 %), community child-care (26 %) and community-welfare centres (28 %). The walking-plus-transit version raises every value but the same five categories remain the planner’s highest priority, becoming the input to the LP relocation.

5.3 LP relocation experiment on the sample sub-region

Component 3 was exercised end-to-end on the sample sub-region (Nam-gu, Busanjin-gu, Dong-gu, Jung-gu, Yeongdo-gu; 820,344 residents) under the target-first operating mode. We ran four LP scenarios in the Living-Zone Dashboard - target-first vs. capacity-gap-min, each in a constrained (planner-curated candidates) and a free (full candidate set) variant - and CBC returned an Optimal status for the primary problem in every case. The experimental setup was as follows: 60 candidate 250 m grid cells (after the planner removed three sites considered politically infeasible from a raw scan of 50 + 13 candidates), a 298-row candidate-type-rule table controlling which facility type is allowed at each candidate site, a fallback shortfall sum of 0.514 (target-first) and a capacity-gap residual of 7.31 across the five optimised types. Of the 60 enabled candidate sites, the LP selected 9 grid cells and co-located 15 new facility units, including a 5-in-1 service node (Mara37ab82ab) that simultaneously hosts a district office, a public-health centre, a sports gymnasium, a community-welfare centre and a retail facility - a configuration that an exhaustive manual search would be unlikely to find. Table 2 reports the per-type coverage change.

Table 2. Walking-only coverage change after LP relocation, sample sub-region (target-first scenario).

Facility type (5 LP-optimised)	Current (%)	After LP (%)	Δ (pp)	Δ population
Public-health centres	52.4	72.1	+19.7	+62,293
Community-welfare centres	58.3	76.0	+17.6	+55,839
Sports gymnasiums	54.9	72.1	+17.2	+54,322
Retail facilities	68.0	82.8	+14.9	+46,981
District offices	57.5	70.1	+12.6	+39,832
All five types (mean)	58.2	74.6	+16.4	+51,853

Figure 4(b) shows the same five rows visually. Three experimental findings stand out. First, the largest absolute gain is at public-health centres (+19.7 pp), which is the type with the second-lowest baseline and the strongest co-location synergy in the LP solution - i.e. the LP exploits the fact that several optimised types can share a single candidate site. Second, even after the LP, no facility type reaches the 90 % target on its own; the residual capacity gap (7.31 across the five types) is something planners can absorb into a longer-term capital plan or close by relaxing the spatial-separation rule. Third, the LP did not consolidate any existing facility (max_remove_count was set to 3 but the LP chose zero removals), which empirically confirms a policy intuition that Choi (2025) sets out for ageing-but-not-yet-shrunken sub-regions: an equity-leaning regime that adds new facilities is more appropriate than a pure efficiency-leaning regime that consolidates existing ones.

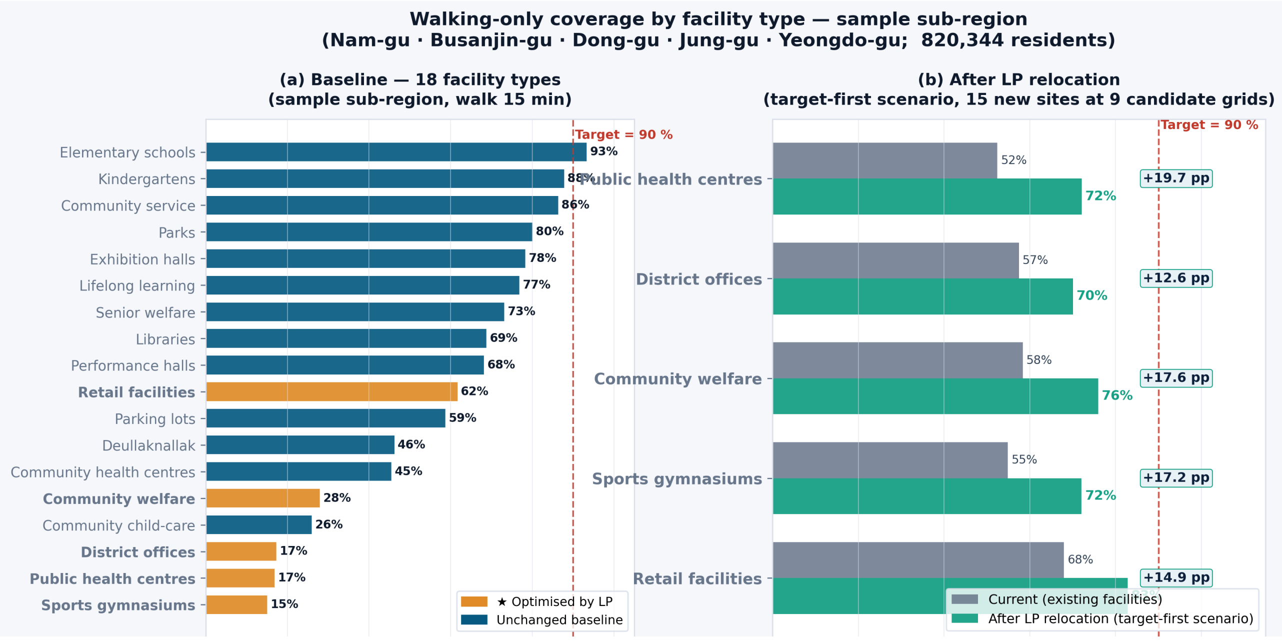

Figure 4. Sample sub-region coverage by facility type. (a) Walking-only baseline across 18 categories - five types fall well below the 90 % target. (b) After LP relocation (target-first, 15 new sites at 9 candidate grids), the five optimised types gain 12–20 pp.

6. Discussion

Three points deserve emphasis. First, the contribution lies in the integration of multiple open geospatial tools into a unified analytical workflow, not in any single component. Each of community detection, multimodal routing and LP optimisation is well-established in isolation, but combining them under one FOSS roof - with consistent intermediate artefacts, deterministic seeds and a single Snakemake DAG - is what makes the framework usable inside a five-yearly Living-Area Plan revision cycle. Second, the choice of α* has substantive downstream consequences. A higher α* yields smaller mobility-derived communities, which restricts the LP feasible space. The α-sweep diagnostic is therefore part of the policy conversation, not an internal technicality. Third, the LP objective’s three λ coefficients are themselves policy choices; the framework returns Pareto frontiers of (cost, accessibility) trade-offs that planners can inspect before committing.

Comparing the framework with current Korean planning practice illuminates the FOSS4G contribution. The 2025 Gangdong Living-Area Plan was drafted using accessibility heat-maps and Louvain zoning but the placement decisions were made by expert deliberation. Our prescription overlaps with the expert proposal, suggesting that the LP captures the same intuition experienced planners apply, with the advantage of being auditable, reproducible and runnable on any city whose data follow the same FOSS schemas.

Several limitations should be acknowledged. The 250 m grid may smooth fine-grained accessibility deserts; a two-resolution toggle is planned. The T-MAP API queries are commercial and not openly redistributable. The equity weight w_i is a policy choice; testing alternatives based on income or elderly share is a priority for follow-up work. The LP assumes that the demand P_i and the candidate set J* are exogenous; a fully dynamic version would couple the LP with a land-use feedback loop. From a FOSS4G perspective, the most pressing extension is to package Component 3 as an OGC API – Processes endpoint so that external GIS clients (QGIS, R, MapLibre browsers) can invoke the optimisation directly.

7. Conclusions

We have presented an open geospatial analytical framework for proximity-based community planning that integrates three components - mobility-based community detection on telecom OD flows, multimodal accessibility analysis with OSRM and r5py on OpenStreetMap and GTFS, and LP facility allocation with a combined equity-and-efficiency objective - into a single reproducible FOSS workflow. Applied to Busan Metropolitan City via the Living-Zone Dashboard package, the framework returns 50 mobility-derived and 50 accessibility-derived communities on a 12,474-cell 250 m grid (Q_γ = 0.934 and 0.918), with a population-weighted IoU of 0.525 (ARI = 0.547, NMI = 0.814) between the two delineations. The LP relocation, exercised on a sample sub-region of five inner-city districts (820,344 residents), places 15 new facility units across 9 candidate 250 m grid cells and lifts walking-only coverage for the five most deficient facility types by 12–20 percentage points each (public-health centres +19.7 pp, community-welfare +17.6 pp, sports gymnasiums +17.2 pp, retail +14.9 pp, district offices +12.6 pp). The entire pipeline runs end-to-end using only free and open-source software.

8. Acknowledgements

This research was supported by Grant RS-2022-00143404 from the Big-Data-Based Artificial-Intelligence Urban-Planning Technology Development Programme funded by the Ministry of Land, Infrastructure and Transport of the Republic of Korea.

9. References

Allam, Z., Bibri, S. E., Chabaud, D., Moreno, C., 2022. The 15-minute city concept can shape a net-zero urban future. Humanities & Social Sciences Communications, 9(1), 126. https://doi.org/10.1057/s41599-022-01145-0

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Choi, J., 2025. AI-based Living Area Planning Support System: In Response to Demographic Structural Change and Proximity-Oriented Daily Living Zones (in Korean). In: Smart City Top Agenda 2025 (KAIA), pp. 217–235.

Dovey, K., Pafka, E., 2019. What is walkability? The urban DMA. Urban Studies, 57(1), 93–108.

Eom, S., 2025. Community-stability assurance for Louvain-based Living-Area zoning (in Korean). Seoul Institute Project Report.

Hagberg, A. A., Schult, D. A., Swart, P. J., 2008. Exploring network structure, dynamics, and function using NetworkX. Proc. 7th Python in Science Conf. (SciPy 2008), Pasadena, CA, USA, 11–15.

Köster, J., Rahmann, S., 2012. Snakemake-a scalable bioinformatics workflow engine. Bioinformatics, 28(19), 2520–2522. https://doi.org/10.1093/bioinformatics/bts480

Lee, Y., Choi, J., Eom, S., Yeo, J., 2025. Data-driven Living-Area planning: a comparative study of accessibility-based and flow-based delineations (in Korean). Proc. Urban Design Inst. Korea Autumn Conf., 243–250.

Luxen, D., Vetter, C., 2011. Real-time routing with OpenStreetMap data. Proc. 19th ACM SIGSPATIAL Int. Conf. on Advances in Geographic Information Systems, 513–516. https://doi.org/10.1145/2093973.2094062

Mitchell, S., O'Sullivan, M., Dunning, I., 2011. PuLP: a linear programming toolkit for Python. The University of Auckland, Auckland, New Zealand.

Moreno, C., Allam, Z., Chabaud, D., Gall, C., Pratlong, F., 2021. Introducing the '15-Minute City': Sustainability, resilience and place identity in future post-pandemic cities. Smart Cities, 4(1), 93–111. https://doi.org/10.3390/smartcities4010006

Park, J., Choi, S., Lim, H., 2023. Big-data-based delineation of 15-minute walkable Living Areas (in Korean). Proc. KSTS–KPA Joint Conf., 638–643.

Pereira, R. H. M., Saraiva, M., Herszenhut, D., Braga, C. K. V., Conway, M. W., 2021. r5r: Rapid Realistic Routing on Multimodal Transport Networks with R5 in R. Findings, 21262. https://doi.org/10.32866/001c.21262

Seoul Institute, 2024. Big-Data-Based AI Urban-Planning Technology Development (4th-year report) (in Korean).

Tenkanen, H., Fink, C., Klumpenhouwer, W., Saraiva, M., Pereira, R. H. M., 2022. r5py: Rapid Realistic Routing with R5 in Python. Zenodo. https://doi.org/10.5281/zenodo.7060437

Wan, N., Zou, B., Sternberg, T., 2012. A three-step floating catchment area method for analysing spatial access to health services. International Journal of Geographical Information Science, 26(6), 1073–1089. https://doi.org/10.1080/13658816.2011.624987

Yeo, J., Lee, M., Han, J., 2025. Multimodal Living-Area delineation for Busan's Gangdong sub-region (in Korean). Gachon University Smart-City Project Report.