Optimal Trajectory Planning of Drones for 3D Mobile Sensing

Anonymous

Abstract—**Projecting the population distribution in geographical regions is important for many applications such as launching marketing campaigns or enhancing the public safety in certain densely-populated areas. Conventional studies require the collection of people's trajectory data through offline means, which is limited in terms of cost and data availability. The wide use of online social network (OSN) apps over smartphones has provided the opportunities of devising a lightweight approach of conducting the study using the online data of smartphone apps. In this paper, we propose the concept of geo-homophily in OSNs to determine how much the data of an OSN can help project the population distribution in a given division of geographical regions. Specifically, we establish a three-layer theoretic framework that first maps the online message diffusion among friends in the OSN to the offline population distribution over a given division of regions via a Dirichlet process, and then projects the floating population across the regions. By experiments over large-scale OSN datasets, we show that the proposed prediction models have a high prediction accuracy in characterizing the process of how the population distribution forms and how the floating population changes over time.**

I. INTRODUCTION

Unmanned aerial vehicle (UAV), commonly known as drone, is an aircraft without a human pilot aboard, which is commonly used in measurement and sampling. Compared to manned aircraft, drones are more suitable for data collections and mobile sensing applications that capture different dimensions of signals in the environment that are beyond our sensing capability, such as aerial photography, 3D wireless signal survey, air quality index (AQI) measurement.

However, civilian drones are still not popular these days. Furthermore, a lot of drone companies were broken down. It could be a quite confusing problem if you have never come into attach with a drone. If you've actually tried using them, you could find that civilian drones do not really apply to daily life due to:

- Low battery available time.
- Great noise during flight.
- Wing rock and more battery drain caused by poor carrying capacity.

Therefore, in order to make more use of existing drones, we must consider the following problem: **How to complete measurement (or flight) in the shortest possible time? Furthermore, in the three-dimensional space?**

Similar to traditional sensor networks and mobile base station, we consider data collection in mobile environment. So total time consumption consists of two parts: **flight time** and **measure time**. While we also have the following difference:

- We consider optimal algorithm in three-dimensional space.
- We use the routing algorithm based on graph theory apart from traditional greedy algorithms.

In this paper, we consider mobile sensing in three-dimensional space. We divide three-dimensional space into a network of observation locations (OLs) and select critical observation locations (COLs) from OLs to cover measurement space, which could be formulated as a constraint set coverage problem in graph theory. Specifically, we consider the following two special cases:

1) *Consider measurement time only*: Under this condition, we assume flight time negligible and consider measurement time only. In order to minimize measurement time, we should select least OLs to cover OL-network. Therefore, we could formulate this problem as a minimum dominating set (A dominating set in a graph G is a subset of vertices S such that every vertex in $V(G)$ is a neighbor of some vertex of S.) problem in grid, which has been studied for a long time.
2) *Consider flight time only*: Under this condition, we assume measurement time negligible and consider flight time only. In order to minimize flight time, we should choose the shortest OL-path in OL-network. Therefore, we could formulate problem as a minimum dominating path (A dominating path is a dominating set as well as a trail where all vertices (except possibly the first and last) are distinct. Briefly, it is a dominating set as well as a path.) problem in grid, which has not been solved before. In this paper, we solved this problem in grid and give an expand in three-dimensional space.

Because of algorithms we use is based on graph theory, We could solve two problems above optimally in O(1)time. We use drones to verify our simulation in multiple scenarios. We find out that the flight time we use is less than ordinary approach.

II. RELATED WORK

A. 3D mobile sensing

B. Route planning in conventional wireless sensor networks

III. SYSTEM MODEL

In this section, we establish a three-dimensional (3D) network model that characterizes the ordinary mobile sensing scene for drone. Then, we analyse correlation between

OLs and relationship between 3D and 2D network model. Afterwards, we formulate the COL selection problem as a constraint set coverage problem. To simplify problem, we consider two special scenes respectively. In the first scene, we only consider measurement time which transform problem into the minimum dominating set problem. In the second scene, we only consider flight time which transform problem into the minimum dominating path problem. We will make further discussion in next subsection. Finally, we define variables that would be used to mathematical proof next section.

A. Network establishment

Dividing a 3D space into cuboids: We assume sensing object is largely same in fixed area for every position in 3D space. So we divide a 3D space into cuboids with a meters long, b meters wide and h meters high. We define the center point of cuboid i as its observation location (OL) (as shown in Figure 4), which is denoted by the 3-tuple (longitude, latitude, and altitude), i.e.,

$$OL_i = (x_i, y_i, z_i),$$

where x_i, y_i, z_i are 3D coordinates of OL_i.

3D network of OLs: The divided cuboids of a 3D space and the corresponding OLs can form a 3D network graph $G = (V; E)$, where V denotes the set of vertices and E represents the edges connecting neighboring vertices. Specifically, the OL inside each cuboid i is considered as a vertex in G, and an edge (i, j) exists if cuboid i is the same as cuboid j in two coordinates and adjacent to cuboid j on the third dimension. Therefore, the 3D network of OLs forms a three-dimensional grid which has fine topology structure.

Levels of OLs: Obviously, cuboids in grid could be classified by height level. We call the ground level as level 1 at height of $\frac{h}{2}m$, one level above as level 2 at height of $\frac{3l}{2}m$.

B. Time consuming

For general mobile sensing, total time consuming consists of flight time and measurement time. And time consuming depends on COLs selected from OLs. We denote V_C as the set of COLs and v_{C_i} as $i-th$ vertex in V_C.

Measurement time: Measurement time T_M is total time spend on mobile sensing. From empirical view, we could assume that measurement time is same for each OL. Therefore, measurement time is proportion to the number of COLs we select. The function is written as

$$T_M = t_M |V_C|,$$

where t_M is the measurement time for each OL.

Flight time: Flight time T_F is total time spend on UAV's flight. Since we formulate 3D space into grid, we use Hamiltonian distance to characterize distance between OLs. So the flight time is proportion to the length of trajectory. The function is written as

$$T_F = t_F \sum_{i=1}^{|V_C|-1} d_H(v_{C_i}, v_{C_{i+1}}),$$

where t_F is the flight time for unit length of coordinate system and d_H denote Hamiltonian distance between two vertexes.

Therefore total time consuming T is

$$T = T_M + T_F.$$

C. Correlation between OLs

To characterize the general mobile sensing process, we assume adjacent OLs have correlation. To characterize different adjacency, we consider following two typical scenarios in mobile sensing.

Star adjacency: In this scenery, we assume OL's neighbors are star adjacent so the sum of three coordinates difference is at most 1 and the coverage set of an OL is the union of its vertex adjacent neighbors and itself. Specifically, an OL has two neighbors in every dimension and the total size of coverage set is 7.

Cubic adjacency: In this scenery, we assume OL's neighbors are cubic adjacent so the max of three coordinates difference is at most 1 and the coverage set of an OL is OLs in a cube whose center is the target. Specifically, an OL has eight neighbors in each plane and the total size of coverage set is 27.

D. Problem formulation

Given a 3D space, we first establish a 3D OL network $G = (V; E)$ which forms a 3D grid. Each OL in grid has a coverage set (that contains correlated OLs with predictable sensing object). Due to drones' limited battery life, we should complete flight and measurement in the shortest time. Hence, we select some OLs as COLs to minimize time consuming while cover whole OL network and formulate the problem as a set coverage problem in 3D grid.

Simplification from adjacency: We have discussed about different adjacencies in the last subsection and in this subsection we can simplify the problem in these scenarios.

1) *Star adjacency*: In actual mobile sensing scenery using UAV, we often consider two dimensions only in this scenery because usually two distant OLs in a line is not predictable, i.e., OLs in different levels. Then, we could divide 3D grid graph into multiple 2D grids and consider set coverage problem in each grid.
2) *Cubic adjacency*: In this scenery, if an UAV flight over a plane, then whole coverage set include the plane and its adjacent planes. Then, we could also simplify 3D grid graph into multiple 2D grids and take advantage of its periodic structure.

Therefore, instead of 3D grid, we formulate the problem into constraint set coverage problem in 2D grid.

Two special cases of time consuming: In section , we have discuss the components of total time consuming. But in actual scene, we usually consider only one part of it. Therefore, in this paper we will consider following two scenarios.

1) *Consider measurement time only*: In this scenery, we assume flight time negligible and consider measurement

time only. As we show in , measurement time is proportion to the number of COLs. So we should select least OLs as COLs to cover OL-network. We could formulate this problem as a minimum dominating set problem in grid.

2) *Consider flight time only*: In this scenery, we assume measurement time negligible and consider flight time only. As we show in, flight time is proportion to the distance between adjacent COLs. Therefore, we could regardless of COL that we take in measurement and select every OL in the path between adjacent COLs where flight time is proportion to the length of UAV's trajectory. So we should find the shortest OL-path in grid which could cover the whole OL-network and select OLs in path as COLs. We could formulate this problem as a minimum dominating path problem in grid.

Therefore, we will discuss these two problems above in the next section and give corresponding certifications.

E. Variable definitions

For the convenience of proof in next section, we define some variables following.

$G = (V, E)$ denotes 3D OL network graph. $L_{m,n}$ denotes grid graph with m rows and n columns. G_i^c and G_i^r the leftmost $i-th$ column and topmost $i-th$ row. For graph G, $V(G)$ denotes set of vertexes in G. And $v_{i,j}$ denotes the vertex in row i and column j in $L_{m,n}$. For any vertex $v, y \in V$, $N[y] = \{v \in V : yv \in E\} \cup \{y\}$ is the closed neighborhood of y (i.e., the set of neighbors of y and y itself). And for $S \subset V$, $N[S] = \bigcup_{v \in S} N[v]$. We denote $G_{i,j}^c$ as columns between G_i^c and G_j^c and $G_{i,j}^r$ as rows between G_i^r and G_j^r. $\gamma(G)$ denotes the domination number of G which is the minimum size of a dominating set of G. $\gamma_c(G)$ denotes the minimum size of a connected dominating set of G. $\gamma_l(G)$ denotes the minimum size of a dominating path of G. For convenience, we assure dominating path L as a special case of connected dominating set which could be represented as a vertex set. We also denote $L_{i,j}^c(G)$ as dominating path $L \cap G_{i,j}^c$ and $L_{i,j}^r(G)$ as dominating path $L \cap G_{i,j}^r$

IV. PROOF

A. Minimum dominating set

B. Minimum dominating path

Lemma 1. *Let $n > 3$, $m > 0$ be integers, and L is a dominating path in $G = L_{m,n}$. Then $|L \bigcap V(G_{n-2,n}^c)| \geq m$. Further, if $3 \nmid m$, then $|L \bigcap V(G_{n-2,n}^c)| \geq m+1$.*

Proof: We follow proof. Since is for connected dominating set and dominating path is a special dominating set, the conclusion as well as the analyzing method of this paper is also applicable to dominating set. ∎

In lemma , we know that every dominating path L in $L_{m,n}$ has at least m vertexes in the three rightmost columns. Therefore, we will consider the three periodicity of dominating path. Specifically, we could construct dominating path in three columns or three rows.

Connecting vertex and dominating vertex: Since dominating path L has both connectivity and dominance, there is thus some vertex v which $N[v] \cap N[L - v] = N[v]$. In other word, the dominating set of v is contained in the dominating set of other vertexes and v is used to connect other vertexes. It is necessary part which connects a dominating set to a dominating path. We defines this kind of vertex as **connecting vertex** whose main effect is connecting vertexes and the other as **dominating vertex** whose main effect is dominating vertexes. Similar with $G_{i,j}^r$ and $G_{i,j}^c$, assume L is a dominating path in G, we denote $D_{i,j}^r(L)$ and $D_{i,j}^c(L)$ as dominating vertexes in $L_{i,j}^r(G)$ and $L_{i,j}^c(G)$ respectively. We also denote $C_{i,j}^r(L)$ and $C_{i,j}^c(L)$ as connecting vertexes in $L_{i,j}^r(G)$ and $L_{i,j}^c(G)$ respectively. And we denote $C_i^{ccon}(L)$ as connecting vertexes between $G_{1,i}^c(L)$ and $G_{i+1,n}^c(L)$ for G with n columns. Besides, we denote $D(G)$ and $C(G)$ as dominating vertexes and connecting vertexes for minimum dominating path of graph G.

Therefore, since we want to prove three periodicity of minimum dominating path in grid, we split dominating path L in $L_{m,n+3}$ into several parts: $L = D_{1,n}^c(L) \cup D_{n+1,n+3}^c(L) \cup C_{1,n}^c(L) \cup C_{n+1,n+3}^c(L) \cup C_n^{ccon}(L)$.

Lemma 2. *Let $n \geq 2$ is integer and L is minimum dominating path of $G = L_{m,n+3}$. We denote $G^* = L_{m,n}$. Then there is at least one condition that $|D_{1,n}^c(L)| \geq |D(G^*)|$. In other words, none of vertexes in $G_{1,n}^c$ has private neighbor in $D_{n+1,n+3}^c(L)$.*

Proof: Since $D(G^*)$ is dominating vertexes in minimum dominating path in $L_{m,n}$ as well as $G_{1,n}^c$ for $L_{m,n+3}$, we have $|D_{1,n}^c(L)| \geq |D(L_{m,n})|$ if all vertexes in $G_{1,n}^c$ is dominated by $D_{1,n}^c(L)$. Therefore, if $|D_{1,n}^c(L)| < |D(G^*)|$, some vertexes in $G_{1,n}^c$ must be dominated by $D_{n+1,n+3}^c(L)$ and do not have neighbors in $D_{1,n}^c(L)$ (as well as $C_{1,n}^c(L)$, but this could make that vertex belongs to $D_{1,n}^c(L)$ instead of $C_{1,n}^c(L)$). Furthermore, these vertexes must belong to G_n^c since vertexes in other columns do not have neighbors in $G_{n+1,n+3}^c$.

$n = 2$: if there is a vertex v in $G_{1,2}^c$ that not be covered by $D_{1,2}^c(L)$ and the row r that v belongs to is not in L. Since v must lie in G_2^c and its left vertex should be dominated by other vertex, the vertex to the up (or down) left of v must belong to L and be one of starting points.

Therefore, we construct dominating path L^* with $|L^*| = |L|$. We replace vertexes in the leftmost column from inflection point to the cross point with G_2^c. As shown in . So $|L^*| = |L|$ and $|D_{1,n}^c(L)| \geq |D(G^*)|$.

$n \geq 3$: Consider there are k continuous vertexes v_k in G_n^c dominated by $D_{n+1,n+3}^c(L)$.

If $k \geq 2$, since these k vertexes are not dominated by $D_{1,n}^c(L)$, the corresponding k vertexes $v_{k'}$ which are in the same row with v_k in G_{n-1}^c can not belong to $D_{1,n}^c(L)$. Therefore, k vertexes $v_{k''}$ in G_{n-2}^c should belong to $D_{1,n}^c(L)$ to dominate $v_{k'}$ because only two endpoints in $v_{k'}$ could be dominated by its top and bottom vertex instead, but their right neighbors could not belong to L which makes L irregular. So we could use the corresponding vertexes in G^c_{n-2} to replace them so as to shorten L. Then, we could use $v_{k''}$ to construct. As shown in .

Since $L^c_{n+1,n+3}$ may have multiple connected components, L may step into $G^c_{n+1,n+3}$ and then move out from $G^c_{n+1,n+3}$ or just step into $G^c_{n+1,n+3}$ and move to the end.

In the first case, Because L may move out from $G^c_{n+1,n+3}$, we could construct as Fig and add connecting vertex to corresponding position.

In the second case, when dominating vertexes in G^c_{n+1} come from G^c_{n+2}, we have the following three cases. When $|v_k| > 3$, L would need more vertexes in G^c_{n+3} to dominate vertexes in G^c_{n+2}. And we could use similar construct like the first case. When $|v_k| < 3$, L will need more connecting vertexes which could also use the same construct. When $|v_k| = 3$ and vertex in v_k do not reach G^r_m, then vertex below $v_{k'}$ must belong to L. So vertexes in Fig is a dominating path for $L_{6,n+3}$ partial but can not reach the minimum so that L could not be the minimum dominating path because the form of minimum dominating path for has same start pointing as L.

When dominating vertexes in G^c_{n+1} do not come from G^c_{n+2}, in other words, L step into $G^c_{n+1,n+3}$ in the first row, move down to $v_{m-1,n+1}$ and use vertexes in G^c_{n+3} to dominate remain vertexes. Since $v_{1,n+1} \in L$, v_k starts from G^r_3. Therefore, we could use similar construct before in Fig to replace L to another dominating path L^* where $|L^*| = |L|$

If $k = 1$, then the vertex must lay in boundary otherwise it will need extra vertexes to connect vertex between G^r_{n+1} and G^r_{n+2}. Therefore, we assume $v_{1,n+1} \in L$. Then $v_{1,n}, v_{1,n-1} \notin L$ and one of $v_{1,n-2}$ and $v_{2,n-1}$ must belong to L to dominate $v_{1,n-1}$. If $v_{1,n-2} \in L$, L will turn to G^r_{n-1} to dominate vertexes in G^r_n and it will bring more vertexes than the following condition. If $v_{2,n-1} \in L$, we will have L like Fig (like $L_{4,11}$). This case could only exist once. We transform $L_{m,n+3}$ symmetrical. Then, $|D^c_{1,n}(L)| \geq |D(L_{m,n})|$ since none of vertexes in $G^c_{1,n}$ has private neighbor in $D^c_{n+1,n+3}(L)$. ■

Lemma 3. *Given L as the minimum dominating path of $G = L_{m,n+3}$, then $|D^c_{n+1,n+3}(L) + C^{ccon}_n(L)| \geq m$. Further, if $3 \nmid m$, then $|D^c_{n+1,n+3}(L) + C^{ccon}_n(L)| \geq m+1$.*

Proof: Since G^c_{n+1} might be dominated by $D^c_{1,n}(L)$, we consider the coverage problem of $G^c_{n+2,n+3}$ only.

Before formal proof, we will prove that expect for one single case, $G^c_{n+2,n+3}$ is dominated by rows. Specifically, every row is dominated by only one connected component in $L^c_{n+1,n+3}$.

If G^r_i in $G^c_{n+2,n+3}$ is dominated by two connected components in $L^c_{n+2,n+3}$, then we assume $v_{i,n+2}$ is dominated by a component above and $v_{i,n+3}$ is dominated by the other component beneath.

Therefore, there are two different scenarios. Under the first scenery, G^r_i is dominated by two end vertexes like Fig which can be transformed by extending one vertex to dominate all vertexes dominated by two components. Under the second scenery, G^r_i is dominated by one end vertex and one intermediate vertex. This is the unique case that could not be replaced. But we could take them as one part since the union of two components follows the result.

Then, we prove the lemma by induction. When $m = 1$, $|D^c_{n+1,n+3}(L) + C^{ccon}_n(L)| \geq 2$. When $m = 2$, $|D^c_{n+1,n+3}(L) + C^{ccon}_n(L)| \geq 3$. When $m = 3$, $|D^c_{n+1,n+3}(L) + C^{ccon}_n(L)| \geq 3$. As shown in .

Now assume the result holds for $m = k$. When $m = k+1$, if there is only one connecting component in $L^c_{n+1,n+3}$, $|\gamma_c(G^c_{n+2,n+3})| \geq m$. Adding 1 connecting vertex in G^c_{n+1}, $|D^c_{n+1,n+3}(L) + C^c con_n(L)| \geq m+1$. If there are multiple connecting components in $L^c_{n+1,n+3}$, we assume $G^c_{n+2,n+3}$ is dominated by rows by proof above. If a rows and b rows are dominated by two connected components L_a and L_b respectively. If $3 \mid m$, then at most two of a and b could be divided by 3 so that $(D^c_{n+1,n+3}(L) + C^c con_n(L)) \cap (L_a \cup L_b) \geq a+b$. If $3 \nmid m$, then at most one of a and b could be divided by 3 so that $(D^c_{n+1,n+3}(L) + C^c con_n(L)) \cap (L_a \cup L_b) \geq a+b+1$. Therefore, multiple connected components can finally reduce to one component which also holds the result. ■

Theorem 4. *Let $m \geq 2$ and $n \geq 2$ as integers. We assume $G = L_{m,n+3}$, $G^* = L_{m,n}$. Then, $\gamma_l(G) \geq \gamma_l(G^*) + m$. Further, when $3 \nmid m$, $\gamma_l(G) \geq \gamma_l(G^*) + m + 1$*

Proof: In lemma and lemma, we know that for there exist at least one minimum dominating path L in $G = L_{m,n+3}$, $|D^c_{n+1,n+3}(L)| + |C^c con_n(L)| + |D^c_{1,n}(L)|$ could fulfill additive part in result. Therefore, if the result if false, $|C^c_{1,n}(L)| + |C^c_{n+1,n+3}(L)|$ in G must be less than $C(G^*)$.

Connectivity on the boundary: Because connectivity depends on structure of G^r_n and there are only two start vertexes in L, we have structures in G^r_n like Fig. We will consider different structures of L in G^r_n.

If there are only one connecting vertex $v_{i,n}$ in connected component, we have following two cases. In the first case, like Fig , we have $v_{i-1,n+1}$, $v_{i-1,n+2}$, $v_{i+1,n+1}$, $v_{i+1,n+2}$, $v_{i,n+2} \in L$. In this case, although $v_{i-1,n+1}, v_{i+1,n+1} \in C^c con_n$ which decreases $|C^c_{1,n}(L)| + |C^c_{n+1,n+3}(L)|$, but $|C^c con_n(L) + D^c_{n+1,n+3}(L)| = 5$ which still use 4 vertexes addition to dominate 3 rows. In the second case, we have $v_{i-1,n+1}$, $v_{i-1,n+2}$, $v_{i-1,n+3}$, $v_{i+1,n+1}$, $v_{i+1,n+2}$, $v_{i+1,n+3}$, $v_{i,n+3} \in L$ where $|C^c_{1,n}(L)| + |C^c_{n+1,n+3}(L)| = |C(G^*)|$.

If there are two connecting vertexes in connected component, when we consider them separately, result is the same as one connecting vertex. If we consider them together, the size of L in G^r_n must be 4 as shown in Fig . We assume two vertexes are $v_{i,n}$ and $v_{i+1,n}$. Also, we have two cases. In the first case, we have $v_{i-1,n+1}$, $v_{i-1,n+2}$, $v_{i-1,n+3}$, $v_{i,n+3}$, $v_{i+1,n+3}$, $v_{i+2,n+1}$, $v_{i+2,n+2}$, $v_{i+2,n+3} \in L$, also $|C^c_{1,n}(L)| + |C^c_{n+1,n+3}(L)| = |C(G^*)|$. In the second case, $|C^c_{1,n}(L)| + |C^c_{n+1,n+3}(L)|$ decreases 2, $|C^{conn}_n|$ add 2 and $|D^c_{n+1,n+3}|$ add 4 which use only 4 vertexes addition to dominate 4 rows. It is less than additive part in result. However, this kind of structure could exist when n is small since to reach that structure, G^r_{i-2} and G^r_{i+3} must be dominated by other components and this destroy the 3 periodic structure which will add more vertexes in dominating path. Specifically, if there are two 2-period extend, the dominated rows will decrease more than gain in the dominating vertexes. And if there are two 3-period extend, it could be replaced by similar structure which move two start points to G^r_{n+3}. Therefore, there are only two possibilities that the structure could exist. In the first possibility, $m \equiv 2 \pmod 3$ and $m \geq 11$, as shown in Fig , we have the structure. Consider the structure in $L_{m',n'}$. We assume the structure first appear when $n' = b$. Since $|L^c_{n'-2,n'}| = m'$, minimum dominating path L' in $n' = b-3$ must hold the same structure which drop $m'+1$ vertexes in origin L'. And this structure will hold till $n' \leq 3$. Except for $n' = 1$ which is out of range, $n' = 2$ or $n' = 3$ do not have the same structure which makes contradictions. In the second possibility, $m' \equiv 1 \pmod 3$. However, we could simplify this case, since when $n' \equiv 0 \pmod 3$ or $n' \equiv 2 \pmod 3$, we could change m' and n'. Therefore, we consider $m' \equiv 1 \pmod 3$ and $n' \equiv 1 \pmod 3$ only. Then, we could simplify to $L_{m',n'}$ to $L_{m',1}$ with the same approach like the first possibility which is also out of range.

Connectivity in the middle: In this case, consider a minimum dominating path L^* in G^* which is the base of L. we split $L^{*c}_{1,n}$ on middle and connect two connected components in $L^c_{n+1,n+3}$ and generate L. We assume the origin endpoints in L^* lay in G^{*c}_n, otherwise we need more connecting vertexes between $L^c_{1,n}$ and $L^c_{n+1,n+3}$. If $C^c_{1,n}(L)| + |C^c_{n+1,n+3}(L)$ could be less than $|C(G^*)|$, we consider $L^c_{1,3}$. In lemma, $|L^c_{1,3}| \geq m$ and if $3 \nmid m$, $|L^c_{1,3}| \geq m+1$. Therefore, we could move connecting vertexes from $C^c_{1,3}$ to $C^c_{4,n}$ if necessary without add more vertexes since there are no start point in G^{*r}_4 and we could reduce the extend from minimum dominating path in $L_{m,n-3}$ to L^*. Then, we could construct a minimum dominating path L^{**} from $L^c_{4,n+3}$ and $|L^{**}| < \gamma_l(G^*)$ and this makes contradiction.

Therefore, since $|C^c_{1,n}(L)| + |C^c_{n+1,n+3}(L)|$ is no less than $C(G^*)$, we could prove the result.

∎

Assume $G = L_{m,n+3}$, we denote **standard extend row** as the minimum dominating path of $L_{3,3}$. This structure is quite useful since it uses only 3 vertexes to dominate 3 rows which is 1 vertex per row. Furthermore, it can extend to the next 3 columns which is scalable. Therefore, with standard extend row, we could extend G while holding minimality of L.

Now, we consider the use of standard extend row in $G^c_{n+1,n+3}$ when $n \geq 4$. We will discuss some cases which $\gamma_l(G) = \gamma_l(G^*) + m$ or $\gamma_l(G) = \gamma_l(G^*) + m + 1$ in theorem .

First, we consider $\gamma_l(G) = \gamma_l(G^*) + m$ when $3 \mid m$. When $n = 3$, $\gamma_l(L_{2,3}) = 2$, otherwise $\gamma_l(L_{2,n}) = n$. Therefore, if there are more than or less than 3 rows in $G^c_{n+1,n+3}$ that dominated by a connected component, we will need one more vertexes to connect just like Fig. In lemma , except for combination of standard extend rows, we could replace any pair connected components with one connected component without the same amount of vertexes. Therefore, to reach $\gamma_l(G) = \gamma_l(G^*) + m$, we need divide each three rows into one group and select the middle one into L which could only realize when $3 \mid m$. Besides, we need the same structure in $L^c_{n+1,n+3}$ as shown in Fig .

Then, we consider $\gamma_l(G) = \gamma_l(G^*)+m+1$. When $3 \mid m+1$, we can also use standard extend row to fulfill the additive part in theorem .

Therefore, when $3 \mid m$ or $3 \mid m+1$, we can use standard extend row to construct L which could reach additive component in the result. And we could combine standard extend row with other structure while hold minimality of vertexes. So we will consider the structure without standard extend rows.

Theorem 5. *We denote s $3-period\ extend$ as extending three columns right from $L_{m,n}$ to $L_{m,n+3}$ where $|L_{m,n+3}| - |L_{m,n}| = s$. Assume $G = L_{m,n}$, except for standard extend rows, when $n \to \infty$, there are no $s 3 period extend$ on average which $s < m+2$. Specifically, there are only finite 3 period extend where vertexes is less than $m+2$ for fixed m.*

Proof: We assume that there are infinite 3-period extends where vertexes is less than $m+2$ for fixed m. As we discuss above, there are only $m+1$ $3-period\ extend$ which is less than $m+2$.

As theorem ,we know that without standard extend row, $|C^{conn}_n + D^c_{n+1,n+3}|$ is more than dominated rows in other structure. And multiple connected components can transform into one connected components so that two start points can merge to one start point. Therefore, we only need to consider the coverage problem with one start point.

Extend from start point: So we will consider 4 different cases of $m+1$ $3-period\ extend$ which is from start point in $L^c_{n+1,n+3}$. As shown in Fig .

For case (a), we need $v_{1,n}$ or $v_{m,n}$ to connect L and case (b) is similar with (a). For case (c), we need corresponding vertexes in G^r_n to dominate uncovered vertexes in G^r_{n+1} and case (d) is the same.

However, case (b) and case (d) can not guarantee connectivity of L for further extends. Under case (b), L do not reach G^r_{n+3} which is still a $m+2$ $3-period\ extend$. Under case (d), except for standard extend row, to extend 3 columns further, L will need two more extra vertexes to connect. Therefore, these 2 cases can only exist in the last 3-period.

Under case (a), since $v_{m-1,n+3} \in L$, to extend 3 rows on right, L will add $v_{m,n+3}$ to reach condition for case (b) which is a $m+1$ $3-period\ extend$. But it will make case (a) as a $m+2$ $3-period\ extend$. Otherwise, holding the $m+1$ $3-period\ extend$ for this period, we will have following 2 cases. To reach $v_{m,n+6}$, we must reach $v_{m,n+5}$ or $v_{m-1,n+3}$ which will add extra vertexes to take a 'detour'. Then, we will need at least $m+3$ vertexes to dominate this period.

Under case (c), we will have similar result. Consider previous period of L. Since G^r_{n+1} should be dominated by corresponding vertexes in G^r_n and due to $3-period$, $v_{1,n-1}$ should be dominated by G^r_{n-2}. Therefore, there are $2m$ vertexes in previous period. So there are at most one case

(c) in L.

Therefore, we have at most one case (c) and at most one case (b) or (d).

Extend from middle: Except for standard extend row, there are only one $s\ 3-period\ extend$ case where $s < m+2$ which is in Theorem. When $n > 1$, it could bring $m\ 3-period\ extend$. But this can only occur once.

Rotate direction: Since we may change the origin structure of L, we could rotate direction of extend direction by 90 degree. Therefore, when $m \equiv 0 \pmod 3$, we may use standard extend row on rotated structure which could bring a $m+1\ 3-period\ extend$. But similarly, this trick could only exist once for the same m.

Therefore, there are only finite $m+1\ 3-period\ extend$ in L.

■

According to Theorem above, we could find out that there are only several cases for $m\ 3-period\ extend$ and $m+1\ 3-period\ extend$, we will show them below.

m 3-period extend: There are only two cases for $m\ 3-period\ extend$. When $m \equiv 0$, $standardextendrow$ could be a $m\ 3-period\ extend$. When $n \equiv 1$, like($L_{10,4}$) could only exist once. When $m \equiv 2$ and $n \equiv 1$, like($L_{11,4}$) could only exist once also.

m+1 3-period extend: In general, We have at most one case (c) and at most one case (b) or (d). We could also use a $m\ 3-period\ extend$ on the L^* where $|L^*| = |L|+1$, i.e., we could rotate the extend rotation by 90 degree and use standard extend row on it. When $m \equiv 2$, like($L_{11,4}$) could only exist once also. When $m \equiv 2$ and $n \equiv 2$, like($L_{5,11}$) could be a m+1 3-period extend. When $m \equiv 0$ and $n \equiv 2$, like($L_{6,11}$) could be a m+1 3-period extend.

We denote $a = \lfloor \frac{m}{3} \rfloor$, $b = \lfloor \frac{n}{3} \rfloor$. And we will show six different cases for minimum dominating path for $L_{m,n}$ and give its proof.

Proposition 1. *When $m \equiv 0 \pmod 3$, $n \equiv 1 \pmod 3$, we have $\gamma_l(L_{m,n}) = 3ab+3a-2$. When $m \equiv 2 \pmod 3$, $n \equiv 1 \pmod 3$, we have*

$$\begin{aligned}&\gamma_l(L_{m,n})\\&= \begin{cases} 3ab+3a+3b & a \le 2 \\ 3ab+3a+3b-1 & otherwise \end{cases}\end{aligned}$$

.

Proof: When $m \equiv 0 \pmod 3$ or $m \equiv 2 \pmod 3$, standard extend row could reach $m\ 3-period\ extend$ and $m+1\ 3-period\ extend$ respectively which is better than other structure.

When $n = 1$, $\gamma_l L_{m,1} = m-2$. L contains $\lceil \frac{m}{3} \rceil$ dominating vertexes and $m-2-\lceil \frac{m}{3} \rceil$ connecting vertexes which can be used to connect standard extend rows. As shown in . Further, when $m > 2$, the structure in could bring a $m\ 3-period\ extend$ when $m \equiv 2 \pmod 3$.

Then, when $m \equiv 0 \pmod 3$, L will add m vertexes in every 3-period while when $m \equiv 0 \pmod 2$, L will add $m+1$ vertexes in every 3-period except for $m = 11$. And for discuss above, when $a > 2$ and $n = 1$, there is a $m\ 3-period\ extend$ for $L_{m,n}$. Therefore due to Theorem they could achieve optimal solutions.

Therefore, when $m \equiv 0 \pmod 3$, $n \equiv 1 \pmod 3$,

$$\begin{aligned}\gamma_l(L_{m,n}) &= a(3b+1)+2(a-1)\\ &= 3ab+3a-2\end{aligned}$$

When $m \equiv 2 \pmod 3$, $n \equiv 1 \pmod 3$, if $a \le 2$,

$$\begin{aligned}\gamma_l(L_{m,n}) &= (a+1)(3b+1)+2(a-1)+1\\ &= 3ab+3a+3b\end{aligned}$$

if $a > 2$,

$$\gamma_l(L_{m,n}) = 3ab+3a+3b-1$$

■

Proposition 2. *When $m \equiv 0 \pmod 3$, $n \equiv 0 \pmod 3$, we have $\gamma_l(L_{m,n}) = \min(3ab+2a-2, 3ab+2b-2)$. When $m \equiv 0 \pmod 3$, $n \equiv 2 \pmod 3$, we have*

$$\begin{aligned}&\gamma_l(L_{m,n})\\&= \begin{cases} min(3ab+4a-2, 3ab+3a+2b-1) & b \le 2 \\ min(3ab+4a-2, 3ab+3a+2b-2) & otherwise \end{cases}\end{aligned}$$

. When $m \equiv 2 \pmod 3$, $n \equiv 2 \pmod 3$, we have

$$\begin{aligned}&\gamma_l(L_{m,n})\\&= \begin{cases} min(3ab+4a+3b+1, 3ab+4b+3a+1) & a \le 2, b \le 2 \\ min(3ab+4a+3b, 3ab+4b+3a+1) & a > 2, b \le 2 \\ min(3ab+4a+3b+1, 3ab+4b+3a) & a \le 2, b > 2 \\ min(3ab+4a+3b+1, 3ab+4b+3a+1) & otherwise \end{cases}\end{aligned}$$

. Besides, the results have a same structure.

Proof: Since we use similar proof methods for 3 cases, we only prove $m \equiv 0 \pmod 3$, $n \equiv 2 \pmod 3$. We will use reduction in a.

When $a = 1$, $\gamma_l(L_{3,n}) = 3b+2$. And the two structures is as shown in Fig ... Since $n \equiv 2 \pmod 3$, we could use $m+1\ 3-period\ extend$ when $n \ge 11$.

Assume we have prove the case when $a \le k$, we consider the case when $a = k+1$. Since we could extend three rows from $a = k$, we could consider the structure of L in $L_{m,n}$. There are also two structures in $L_{m,n}$. We consider critical point between $3ab+4a-2$ and $3ab+3a+2b-1$ or $3ab+3a+2b-2$.

When critical point is between $3ab+4a-2$ and $3ab+3a+2b-2$, when $b \le \lfloor \frac{k}{2} \rfloor$, $\gamma_l(L_{m,n}) = 3ab+3a+2b-2$ which has structure (a), we could use standard extend row to reach a $n+1\ 3-period\ extend$ which is the lower bound. When $b > \lfloor \frac{k}{2} \rfloor$, $\gamma_l(L_{m,n}) = 3ab+4a-2$, and we could only have a $n+2\ 3-period\ extend$ from its structure. L in $L_{3k+3,3b+2}$ is $3kb+3k+5b+1$ and L in $L_{3k+3,3b+5}$ is $3kb+7k+3b+5$. $3kb+3k+5b+1+3k+5 = 3kb+6k+5b+6 \ge 3kb+7k+3b+5$. Therefore, L in $L_{3k+3,3b+2}$ is a minimum dominating path and there are no $m+13periodextend$ or $m\ 3-period\ extend$ for this structure, L in $L_{3k+3,3b+5}$ is also the minimum dominating path with structure (b). Therefore, we could use standard extend rows in structure (b) to reach optimal solution. When critical point is between $3ab+4a-2$ and $3ab+3a+2b-1$ result is the same.

We could use reduction to prove the result on the other two cases. When $m \equiv 0 \pmod 3$, $n \equiv 0 \pmod 3$, we could use $m+2\ 3-period\ extend$ on $L_{3k,3k-3}$ to get the optimal result on $L_{3k,3k}$. Similarly, we could use similar approach when $m \equiv 2 \pmod 3$, $n \equiv 2 \pmod 3$.

■

Proposition 3. *When $m \equiv 1 \pmod 3$, $n \equiv 1 \pmod 3$, $a \geq 1$ and $b \geq 1$, we have*

$$\gamma_l(L_{m,n}) = \begin{cases} 3ab+3a+3b-3 & a+b \leq 4 \\ 3ab+2a+2b+1 & otherwise \end{cases}$$

.

Proof: Consider the condition when $a=1$. When $b=1$, we have $\gamma_l(L_{4,4})=8$ with following structure. Therefore, we could use $m+1\ 3-period\ extend$ case (b) and case (c) respectively to get $\gamma_l(L_{4,7})=13$ and $\gamma_l(L_{4,10})=18$. Tail now, $L_{4,10}$ is a $m\ 3-period\ extend$ for $L_{1,10}$ using ... and reach the lower bound on Theorem. When $b \geq 4$, we could also get $m\ 3-period\ extend$ on $L_{4,n}$ using standard extend row on the extend of $L_{4,10}$. Therefore, when $b \geq 3$, we have $\gamma_l(L_{4,3b+1})=6b$.

When $a=2$ and $b=2$, we could use a $m+1\ 3-period\ extend$ on $L_{4,7}$ and get $\gamma_l(L_{7,7})=21$

When $a \geq 2$, we know that case (b), case (d) and structure ..could only exist once. Therefore, when $b \geq 3$, we have $\gamma_l(L_{3a+1,3b+1})=3ab+3a+3b-3$ which is $a-1$ times $m+2\ 3-period\ extend$ from $L_{4,3b+1}$.

Therefore, when $a+b \leq 4$, we could use a $m+1\ 3-period\ extend$ for each add one on a or b on the basis of $L_{4,4}$.

■

V. Optimal Trajectory Planning Algorithms

VI. Evaluation

VII. Conclusions

In this paper, we propose a systematic study on the population distribution projection over offline geographical regions by analyzing the geographical attributes of online social networks (OSNs). We propose the concept of geo-homophily in OSNs to establish the correlation between online message diffusion and the stability of geographical regions where a population distribution can be drawn. We formulate the population distribution problem from the perspective of Dirichlet process, and present prediction models to show the process that OSN users are distributed into regions, and infer the floating population across regions. By experiments over the large scale datasets, it is shown that the online message diffusions can help evaluate the stability of geographical regions, which further facilitates the determination of population distribution over fixed regions; the proposed prediction models have a high prediction accuracy in inferring the change of floating population across regions.

References