A Potential Application of k-means Clustering to Base Positioning of First Responder Police Surveillance UAVs

---

Introduction

FAA regulations (FAA 2016) on the use of non-military use of unmanned aerial vehicles (UAVs) currently prohibit flying beyond visual line of sight (VLOS) of the human pilot in command, flying over people, and flying during non-daytime hours. Development of smart (computer navigated) UAV fleets which support police response therefore remains largely futuristic. However, the FAA s Unmanned Aircraft System (UAS) Integration Pilot Program (IPP) has given a select group of research teams the ability to relax regulations such as VLOS. The Chula Vista (San Diego County) police department is one of the IPP awardees and has already demonstrated the value of UAVs in supporting police response in its Drones and First Responder program (Chula Vista Police, n.d.). In particular, since surveillance UAVs can arrive at the crime scene ahead of squad cars, police commanders can assess the situation in advance and thus make better informed decisions of the type of response which is appropriate.

Chicago is the third largest city in the U.S., with extremely high homicide rates in several communities on its west and south sides. This has been linked to unusually high levels of illegal guns and gang activity. The city is divided into 25 police districts, with each district further subdivided into police beats. Public outcry after the brutal 2014 fatal shooting of 17 year old LaQuan McDonald by a police officer resulted in a federal investigation revealing a serious deficiency in the training of police officers, especially the use of force.

A question arises whether a smart-navigated UAV surveillance fleet as being tested in Chula Vista might help to reduce tragedies of excessive use of force in Chicago. To this end, we use k-means clustering based on past homicide data to recommend where a dozen first responder surveillance UAV bases might be positioned as a possible violence reduction measure should smart navigated UAV systems one day be accepted as a best practice in law enforcement.

Data

Chicago homicide data can be obtained from the city of Chicago's Data portal. For the purpose of this lab, we will wrangle the data to use just the lat-lon locations of the homicides. Note that we must eliminate data where the location is unknown.

	ID	Case Number	Date	Block	IUCR	Primary Type	Description	Location Description	Arrest	Domestic	Beat	Ward	FBI Code	X Coordinate	Y Coordinate	Year	Updated On	Latitude	Longitude	Location
0	24787	HN623995	09/20/2007 06:52:00 PM	015XX S SANGAMON ST	110	HOMICIDE	FIRST DEGREE MURDER	APARTMENT	True	False	1232	11.0	01A	NaN	NaN	2007	10/10/2019 04:20:40 PM	NaN	NaN	NaN
1	21056	HW431623	08/31/2013 11:19:00 AM	117XX S HALE AVE	110	HOMICIDE	FIRST DEGREE MURDER	YARD	True	False	2212	34.0	01A	1164670.0	1826809.0	2013	10/10/2019 04:11:30 PM	41.680366	-87.672864	(41.680366299, -87.672864196)
2	1710	HH609795	08/28/2002 08:30:00 AM	030XX S WABASH AVE	110	HOMICIDE	FIRST DEGREE MURDER	SCHOOL YARD	True	False	133	3.0	01A	1177222.0	1884824.0	2002	10/10/2019 04:11:30 PM	41.839293	-87.625170	(41.839292845, -87.625169769)
3	24780	JC449748	10/02/2019 02:56:00 PM	030XX S ST LOUIS AVE	110	HOMICIDE	FIRST DEGREE MURDER	AUTO	False	False	1032	22.0	01A	1153517.0	1884183.0	2019	10/09/2019 04:25:23 PM	41.838037	-87.712173	(41.838037246, -87.712173487)
4	24776	JC456756	10/02/2019 09:15:00 AM	018XX S HAMLIN AVE	110	HOMICIDE	FIRST DEGREE MURDER	STREET	False	False	1014	24.0	01A	1151345.0	1890735.0	2019	10/09/2019 04:25:23 PM	41.856060	-87.719972	(41.856059573, -87.719971925)

Figure 1: Chicago homicide data.

	Latitude	Longitude
1	41.680366	-87.672864
2	41.839293	-87.625170
3	41.838037	-87.712173
4	41.856060	-87.719972
5	42.021638	-87.670717

Figure 2: Chicago homicide data after wrangling.

Results

By using a machine learning method known as k-means clustering, based on the historical record of homicides obtained from Chicago's Data Portal, we are able to locate the lat and lon for 12 base positions (Table 1 and Figure 3). The counts can then be used to determine how many UAVs to position at each base (for simplicity 1 or 2). For example, since Base 1 has the highest count of 1348, we would position 2 UAVs at that base. Base 6, on the other hand, has a low count of 192, so we would only position 1 UAV at that base.

Note that k-means clustering is an intrinsic classification method which uses only the position of points to form the clusters. In particular, it does not take into consideration external information such as police districts or police beats. Figure 4 takes data found in a geojson file to overlay police beats on the k-means clustering map. This helps us to see the prevalence of homicides on the west and south sides of the city.

Project Files

A Github repository (https://github.com/pisihara/kMeansClusteringChicagoHomicides) contains the Jupyter Notebook used to produce Figures 1-4.

References

Chula Vista Police, n.d. UAS Drone Program. Available at https://www.chulavistaca.gov/departments/police-department/programs/uas-drone-program.

Federal Aviation Administration 2016. Fact Sheet – Small Unmanned Aircraft Regulations (Part 107) Availabe at https://www.faa. gov/news/fact_sheets/news_story.cfm?newsId=20516

BASE	Latitude	Longitude	COUNT
1	41.804048	-87.635544	1348
2	36.619446	-91.686566	1039
3	41.901364	-87.760368	874
4	41.690041	-87.629538	1011
5	41.911017	-87.715490	921
6	41.740991	-87.562577	192
7	41.791147	-87.698265	1037
8	41.980917	-87.679242	817
9	41.864763	-87.717392	717
10	41.757314	-87.657711	452
11	41.877233	-87.656014	562
12	41.759845	-87.605234	839

Table 1: Base positions and homicide counts.

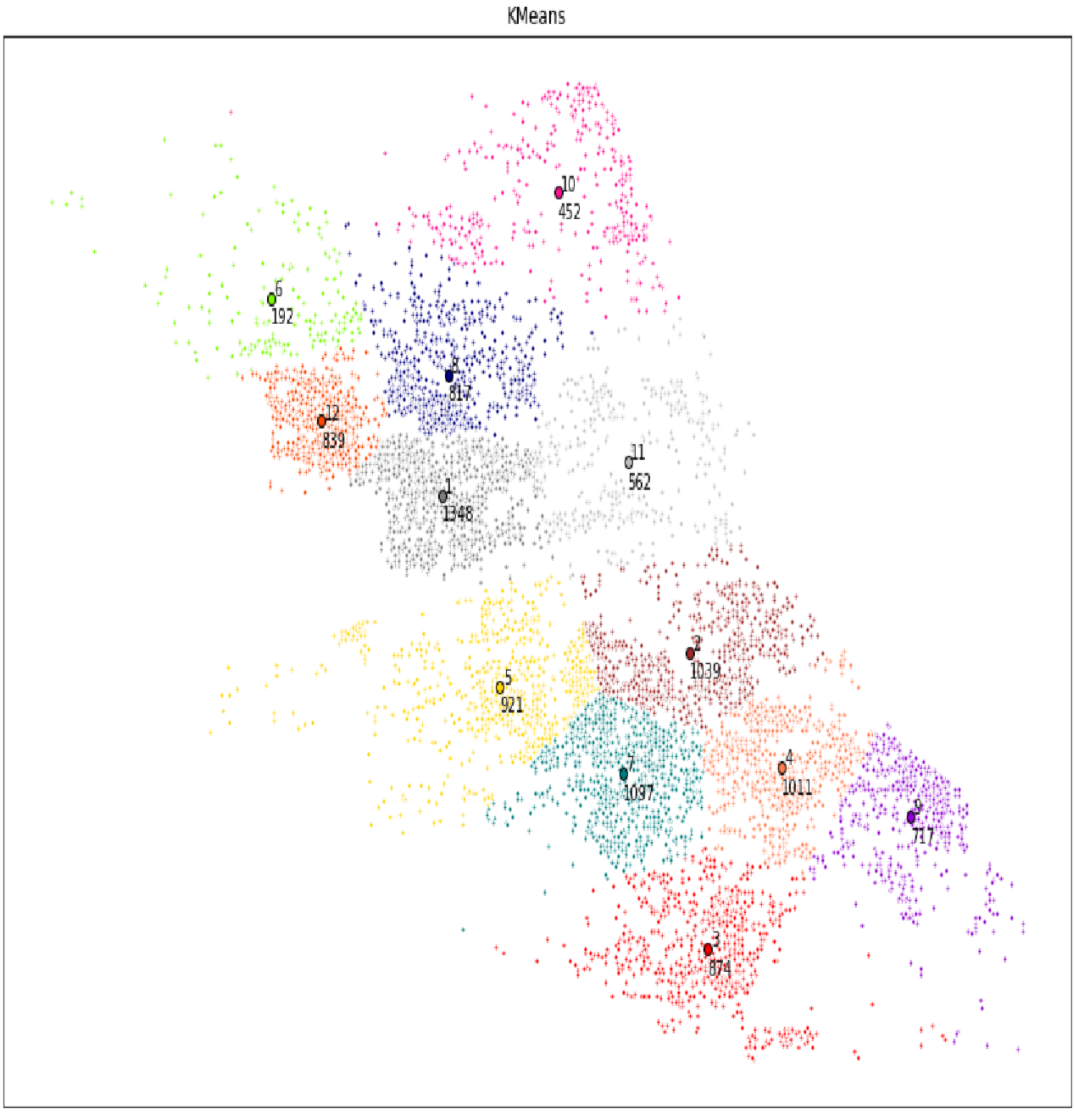

Figure 3: Geographic locations of the base positions and homicides.

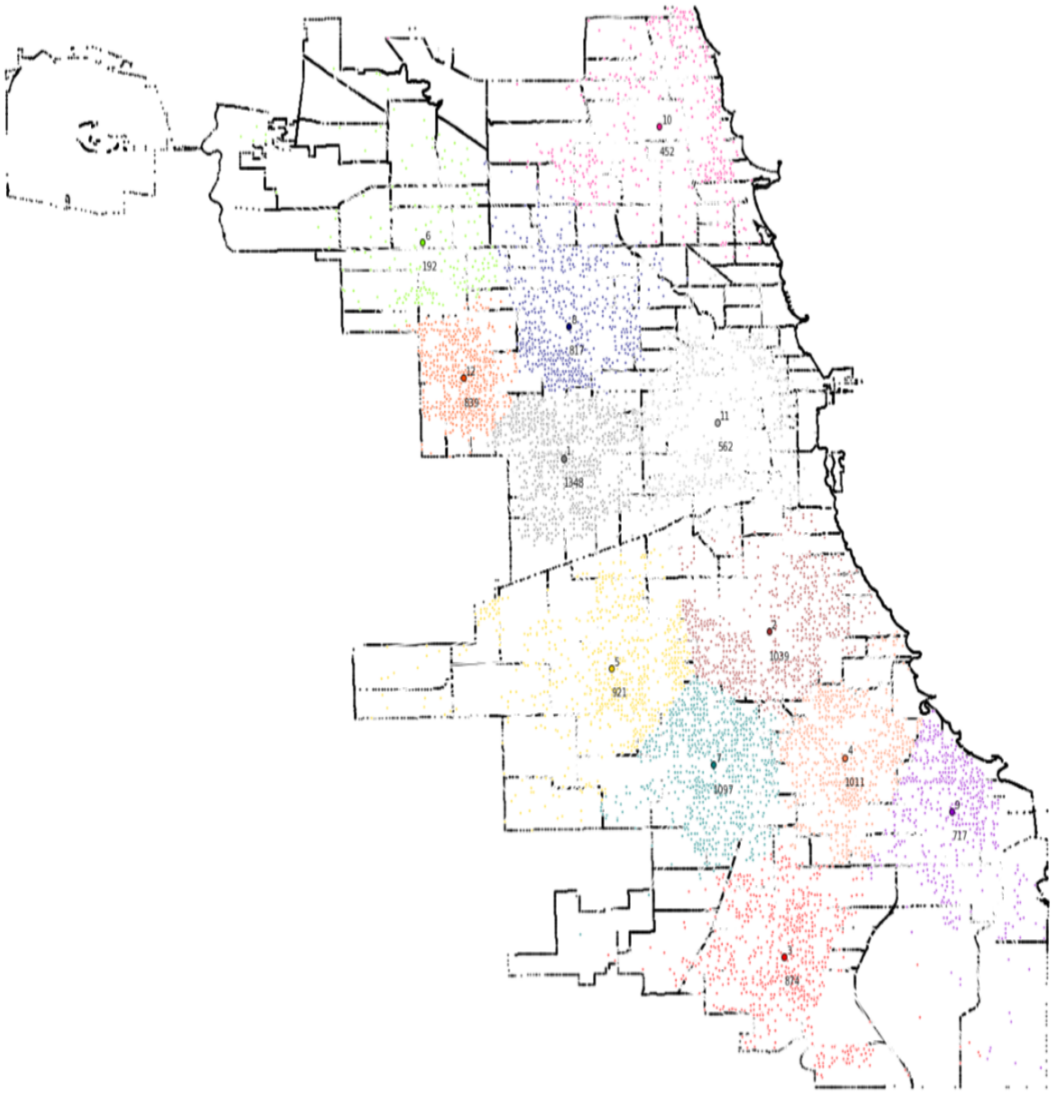

Figure 4: Overlay of police beat boundaries shows heavy clustering of homicides on Chicago's west and south sides. O-Hare airport is located at the top north west corner of the map.