**Anticipation of Locations For Attack**

**(ALFA)**

When a serial criminal targets a city with violent crime, and the police have little to no information about the criminal, ALFA, can predict where the criminal will strike next, with an effectiveness of up to 80%. We devised two schemes compiling different sets of data, in order to predict both the area where the criminal is likely to strike, with an accuracy of 50-75%, as well as the type of victim most likely to be chosen, with an accuracy of up to 75%. The first scheme is based on the geographic locations of the previous crimes. Once these locations are known, analysts draw a series of circles to determine where the highest density of attacks is. The distance between the furthest two points serves as the diameter of the first circle. Then an algorithm is given to draw successively smaller circles, until the closest two points are connected. The area enclosed by the most circles is where the criminal is most likely to attack. Our second scheme looks as the demographics of the previous victims, and uses a series of equations to determine Sx,y, to weight each of the relevant factors, and then determine what area of the city has population demographics most similar to the demographics of the victims, and thus where an attack is most likely to take place.

After both of these areas are calculated, an overlay of the city is created for each scheme, with the areas most likely to be attacked made darkest, and least likely to be attacked made the lightest. These overlays are then placed on top of each other, and the area is the darkest overall is where the criminal is most likely to strike next. ALFA provides a great deal of accuracy because we found a way to organize all of the relevant factors in a series of seemingly random attacks in a way that is both logical and efficient. Since most of the demographical data can be compile far in advance, the implementation of our method will take very little time, thus allowing law enforcement to go out and catch the criminal before someone else is harmed.

**Introduction**

Serial killers perfectly embody the terror of random violence inflicted on friends and loved ones. Their attacks seem to have little to no pattern behind them, but at the same time there is the inherent feeling that there is some meaning behind what is happening. Our algorithm attempts to find the best model for what is essentially unpredictable violence. In this solution, we define an effective algorithm as one that most accurately predicts attacks based on prior location and demographic information.

By examining the previous patterns of the criminal to determine where he or she is comfortable committing crime, we developed a way to approximate the logical places for the next attack. Combining the demographic information of previous victims with knowledge about the demographic information of the surrounding area, we found a way to see where the killer will have the most access to his or her target victims. Using both of these algorithms together, we determined where the criminal is both most comfortable attacking and where it is most convenient for them to do so. This algorithm provides law enforcement with a way to organize random crime in order to prevent future attacks.

**Assumptions**

* **“Serial Crime” is defined as serial murder, rape, and arson.**
* **The criminal has committed a crime at least 5 times.**
* **We know nothing about the criminal.**  Age, psychological history, ect., are unknown. If there was additional data, psychological profiling or another method may be more effective.
* **The residence, or home base, of the killer is unknown.**
* **The killer tends to operate closer and closer to home.** As the killer becomes more experienced, their confidence increases, and they do not feel the need to go as far from home base (investigativepsych.com/geographicprofiling.pdf).
* **The bodies of all previous victims have been found.** Even if the bodies in their entirety have not been found, at least a portion of the victim has been located, in order to gather data for geographic profiling.
* **All previous crimes occur in the same general vicinity.** We are assuming that the criminal is not operating in several states, or across country.
* **There is no blatantly obvious pattern to the criminal’s events.** If the criminal were striking in an exceptionally obvious pattern, like a circle or a cross, then clearly law enforcement should just follow the pattern.
* **The criminal does not attack in the same place twice.** This is similar to attacking in a predictable pattern.
* **We have access to population statistics for the area where the criminal is attacking.**

**Variables**

* **Sx,y -**Likelihood of the criminal attacking within the square (x,y).
* **PD** –The total population representing the demographics of the killer’s previously targeted victims in each square.
* **d**- Distance of the square from the determined center of mass of all incidents and evidence.
* **CM**- Center of mass of all incidents and evidence.
* **v**- The total number of victims
* **vi –**The number of victims representing a given demographic, where the subscript determines the demographic. *m* represents males. *f* represents females. *lc* represents a lower class socioeconomic background. *mc* represents a middle class background, and *uc* represents an upper class background. *o* represents individuals older than 50, *a* represents individuals between 20 and 50 years of age, and *c* represents individuals younger than 20.
* **φi** –The ratio of the number of victims with certain characteristics compared to the total number of victims, with i representing the same demographics as above.
* **pi –**The population of each demographic within a square, with i representing the same demographics as above.
* **Tp-**The total population of a given square of the grid.

**Method**

**Scheme 1**

Our first scheme is recommended for use after compiling information about a serial criminal for at least five different events in order to ensure the greatest accuracy. Conceptually, we wish to find a locus about which the criminal has acted. Based on our assumption that the criminal tends to more closer and closer to his or her “home base,” as he or she becomes more confident, we can attempt to predict the area within which the criminal will act next. We developed an algorithm based on drawing smaller and smaller circles around points where the criminal has attacked, in order to determine where the densest portion of attacks occurred. From this we can narrow down the area in which the criminal is likely to attack next, based on the density.

In order to expand the area around which the criminal will act, we first span the area by drawing a circle with a diameter created by the two events with the greatest distance between each other. This circle will contain most of the other events. Now we attempt to find the more common areas where the criminal will act by creating more circles using the locations of prior events. The next circle on the map will have a diameter the distance of the two events with the second greatest separation. This circle is omitted if one of the two events was used in the prior circle and a circle is drawn with the next two furthest crime points no matter their association with the prior circle. This process is then repeated until the two events with the least distance have been used as a circle or have been skipped due to sharing a point with the prior circle. Each circle that is created has a similar shade of darkness and the more circles that cover a certain area the higher the probability of the serial criminal acting in this area. This will produce a type of density graph that shows the areas within which the criminal is most comfortable. By omitting the points shared with prior circles that would have been the next greatest diameter we omit the random areas of the graph and make the density higher in the areas where more events occurred.

See Image 6 for a visual example of this algorithm.

**Scheme 2**

This scheme is based on our knowledge of the demographics of the previous victims, as well as the population demographics of the city or area where the killer is operating. By looking at the ratios of these demographics, we can attempt to see where it is most practical for the criminal to attack next, assuming his targeting pattern will continue to mimic his previous victims, at least to some extent. We attempted to pick relevant, and available demographic information applicable in many cases. Obviously, if a criminal is targeting a certain, more specific group of victims, such as CEO’s or clergy, that will become a new, and very important, factor. The demographics we chose to include are:

* Gender (male or female)
* Age (child, adult, or elderly)
* Socioeconomic Stats (upper, middle, or lower class)

We tried two variations of this scheme in order to find the one with the greatest accuracy.

**Variation 1**

First we divide the applicable area into grids of a reasonable size. Then we calculate the likelihood of the criminal attacking within that grid, as follows.

Given the total number of victims, *v*, we calculate the ratio of each demographic of the victims, for example male victims (vm), as follows,

φm=vm/v.

We repeat this process for each of our chosen demographics to obtain,

φm, φf, φlc, φmc, φuc, φc, φa, φo

We also assumed that we know the population for each of our grids, *Tp*. We use this to calculate the population density, *PD*, for each grid, which determines which grids have the highest populations that the criminal is targeting. We calculate *PD* as follows,

PD=(1/Tp)(pmφm + pfφf + plcφlc + pmcφmc + pucφuc + pcφc + paφa + poφo).

Now we calculate the center of mass (CM) of all of the instances related to the case, and the distance (d) of the center of each grid, from the center of mass. Using *d*, we calculate,

Sx,y=PD/d.

Now we color in each of the squares of the grid based on the likelihood of an attack in that square. The higher the probability of an attack, the darker the square will be.

**Variation 2**

This variation is essentially the same as the first variation, except when calculating *PD*, we do not take into account the total population of each square of the grid. Thus we have,

PD=(pmφm + pfφf + plcφlc + pmcφmc + pucφuc + pcφc + paφa + poφo).

We compare the differences in the outcomes of these variations in our example below.

**The Method**

Neither of these schemes alone takes into account enough factors to accurately determine where the criminal’s next strike will be. However, by combining our two schemes, we can take into account both where the criminal is comfortable committing crimes, and where victims who are similar to those being targeted can be found.

So first we overlay a map of the region with the density that we found using Scheme 1, to see where the criminal is most comfortable. Then we overlay that with the density that we found based on the demographics of each square, using Scheme 2. Combining these densities, we can see that the darkest areas are the areas that are both near where the criminal feels comfortable, and where the population most strongly resembles his or her target demographics. Also the most likely time of the

**Example**

In order to demonstrate the particulars of our method, we devised our own simple example looking only at the gender that the killer was targeting. In this case 4 males and a female were killed (Table 1). Using completely random values for all variables, we first calculated m= 0.8 and f =0.2 by looking at the criminal’s previous victims. Next, by using the previous crime locations as well as the locations of all the evidence provided (Table 2), we were able to calculate the center of mass (CM) to be (2,2.5). We then used the CM, as well as our information about the population demographics (Table 3) to devise the PD both with and without Tp. (CM is denoted by a star (\*) on all graphs.) The last calculations made before graphing was Sx,y. This was found by using the distance between the midpoint of each grid and then dividing PD (Table 4) and PD(1/Tp) (Table 5) by this distance.

Once all of these calculations were done we used our Scheme 1 method to generate our first graph and found the first and second most dense regions (Image 6). Next, using Scheme 2, we graphed the Sx,y for both PD (Image 7) and PD(1/Tp) (Image 8). Typically these graphs would be shaded darker where the Sx,y is larger but, we found it simpler to write the numbers in the girds. When looking Scheme 1 and both Scheme 2’s we can see that the Scheme 2 from PD(1/Tp) coordinated better with Scheme 1. From this we decided to use PD(1/Tp) to generate our final result..

      In the end we took the top two densest areas for both Scheme 1 and Scheme 2 (from PD (1/Tp)) to get our last graph (Image 9). The darker shading is where they over lapped exactly and is our more accurate guess for the next crime scene location and the other shading would be our next best guess.

**Table 1:** Previous Crimes

|  |  |  |
| --- | --- | --- |
| Victim | Place | Time |
| 30 year old female | (1.5,3) | 1:30am |
| 25 year old male | (3,3.5) | 12:30am |
| 15 year old male | (4,1.5) | 11:00pm |
| 20 year old male | (2,4) | 1:00am |
| 19 year old male | (1,2) | 11:30am |

**Table 2:** Evidence

|  |  |
| --- | --- |
| Type of Evidence | Location of Evidence |
| Body | (1,2.5) |
| Body | (2,3) |
| Letter | (4,2) |
| Letter | (2,2) |

**Table 3:** Total Population: 10,000

|  |  |  |
| --- | --- | --- |
| Approximate Population per grid: | Approximate number of females per grid: | Approximate number of males per grid: |
| (1,1)=400  (1,2)=700  (1,3)=800  (1,4)=600  (2,1)=650  (2,2)=375  (2,3)=500  (2,4)=550  (3,1)=800  (3,2)=600  (3,3)=700  (3,4)=350  (4,1)=675  (4,2)=500  (4,3)=525  (4,4)=375 | (1,1)=250  (1,2)=400  (1,3)=400  (1,4)=350  (2,1)=300  (2,2)=175  (2,3)=200  (2,4)=200  (3,1)=200  (3,2)=500  (3,3)=200  (3,4)=150  (4,1)=275  (4,2)=200  (4,3)=225  (4,4)=250 | (1,1)=150  (1,2)=300  (1,3)=400  (1,4)=250  (2,1)=350  (2,2)=200  (2,3)=300  (2,4)=350  (3,1)=600  (3,2)=100  (3,3)=500  (3,4)=200  (4,1)=400  (4,2)=300  (4,3)=300  (4,4)=125 |

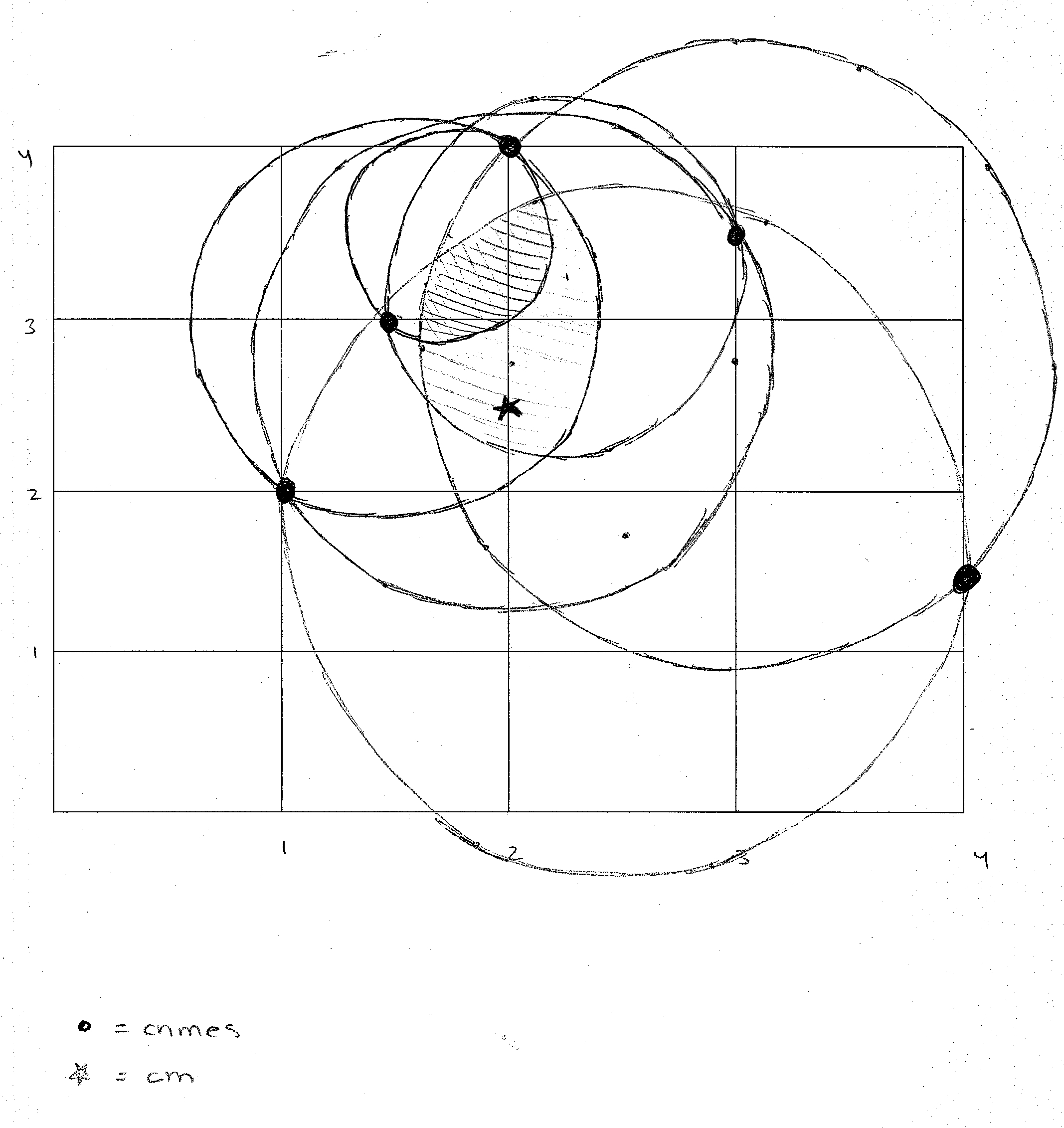
**Table 4**

|  |  |  |
| --- | --- | --- |
| Grid | PD | Sxy |
| (1,1) | 170 | 425 |
| (1,2) | 320 | 155 |
| (1,3) | 400 | 194 |
| (1,4) | 270 | 108 |
| (2,1) | 340 | 227 |
| (2,2) | 195 | 130 |
| (2,3) | 280 | 187 |
| (2,4) | 320 | 178 |
| (3,1) | 520 | 347 |
| (3,2) | 180 | 360 |
| (3,3) | 440 | 880 |
| (3,4) | 190 | 127 |
| (4,1) | 375 | 208 |
| (4,2) | 280 | 187 |
| (4,3) | 285 | 190 |
| (4,4) | 150 | 83 |

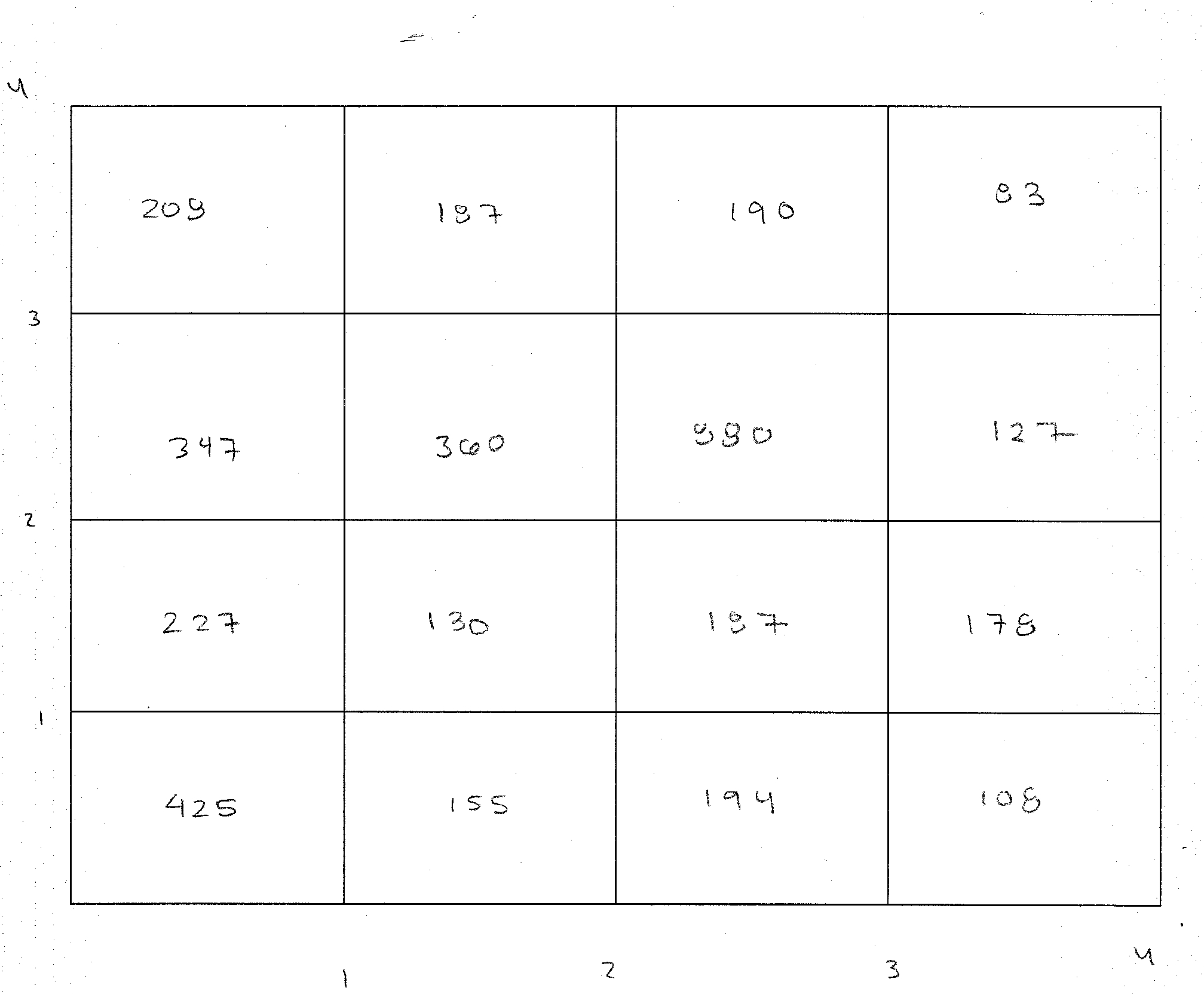
**Table 5**

|  |  |  |
| --- | --- | --- |
| Grid | PD (1/Tp) | Sxy |
| (1,1) | .425 | .170 |
| (1,2) | .457 | .222 |
| (1,3) | .500 | .242 |
| (1,4) | .450 | .180 |
| (2,1) | .523 | .290 |
| (2,2) | .520 | .347 |
| (2,3) | .560 | .373 |
| (2,4) | .582 | .323 |
| (3,1) | .650 | .433 |
| (3,2) | .300 | .600 |
| (3,3) | .629 | 1.258 |
| (3,4) | .543 | .362 |
| (4,1) | .556 | .301 |
| (4,2) | .560 | .337 |
| (4,3) | .543 | .362 |
| (4,4) | .400 | .222 |

**Image 6**



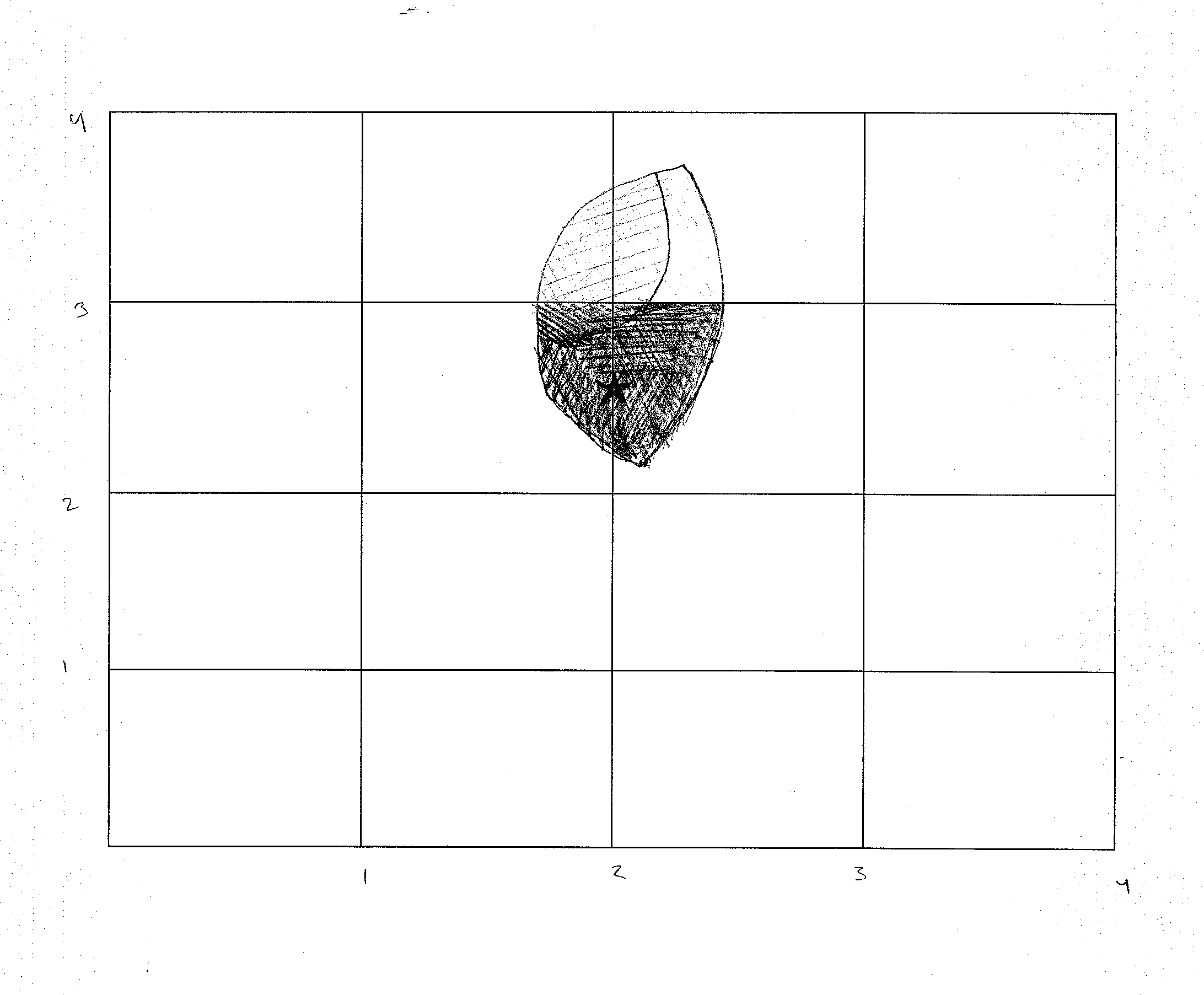
**Image 7:** Sxy Using PD



**Image 8:** Sxy Using PD(1/Tp)



**Image 9:**



**Reliability and Warnings**

This model is most effective when the amount of data increases. With little data neither scheme is effective, with any less than 5 sufficient and distinct pieces of evidence the model is most likely less than 25% effective. This is due to the fact that each data box will have similar density probabilities for suspects. When the distinct amounts of data raises above 10, if a pattern is discernable among victims, the scheme will probably be able to give two or three areas that are denser than the rest of the graph.  Thus narrowing the search to only a few areas and the model is estimated to be around 50-75% effective in a city or isolated area. The mix of the two schemes will be more effective by around 5%, because distance is already taken into account by the second scheme. There is a direct positive correlation between the quantity of solid evidence with discernable patterns and the reliability of this model.

The first scheme alone as we have shown and as has been stated in other works is not so effective. In the cases where the suspect continues to get more comfortable and moves closer to their home this may be as much as 50% or 75% effective. On the other hand, in a case like the BTK case (Appendix A), where the suspect was uncomfortable and continued to expand their area of attack this method would not be effective. It would be close to 10% or less effective. It depends heavily on the case and is based only in the location of events, which may be entirely random. In the case for example where only schools are attacked, clearly if no schools reside in the circle covered areas then this model is not going to be reliable.

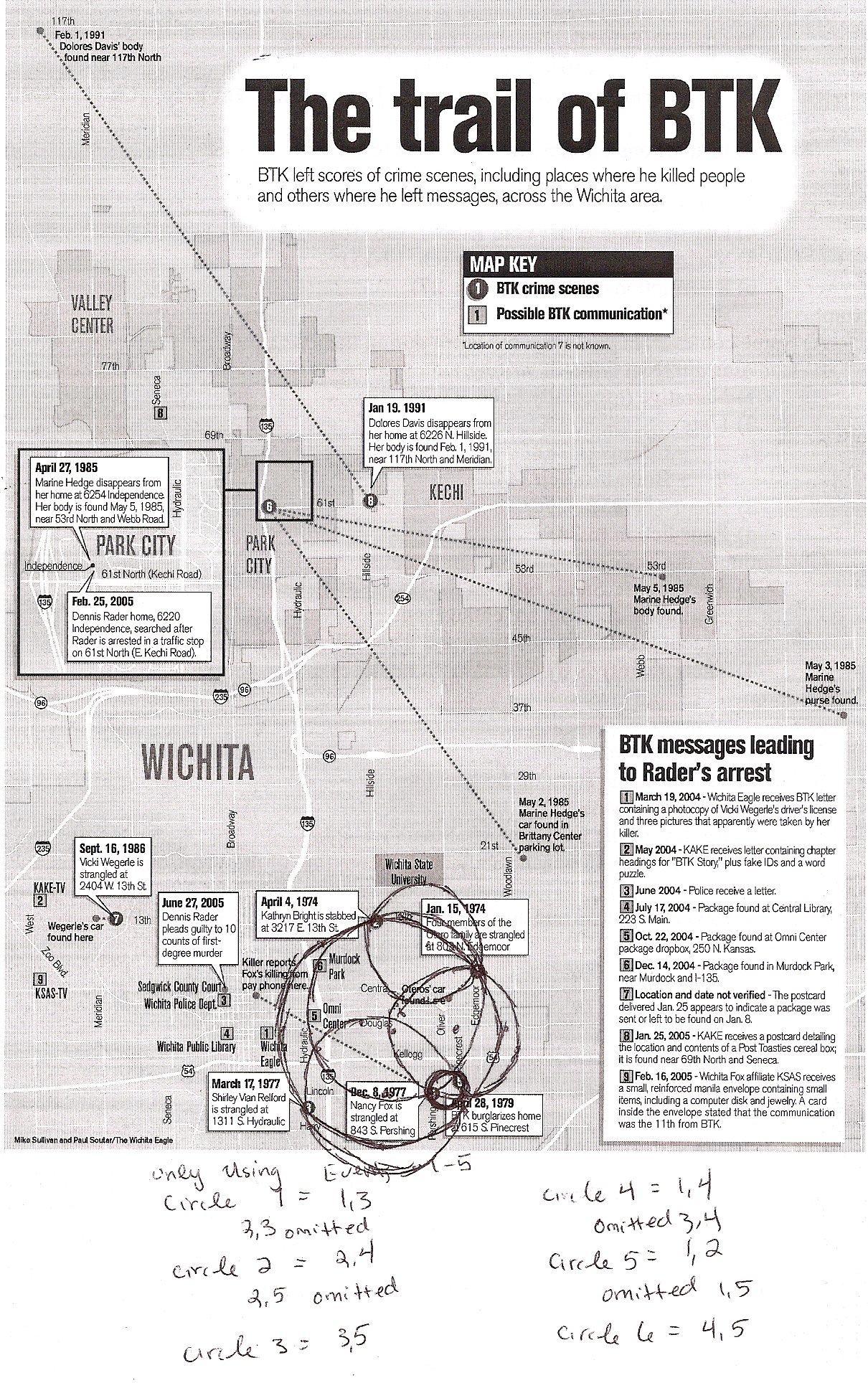
**Conclusion**

One of the strengths of our method is that we do not attempt to find a hidden meaning or explanation in the actions of a deranged killer. Instead we rely on the facts of their past actions, as well as some common traits of in the behavior of random crime, to attempt to discern where the next attack will take place. By taking information from both the criminal and the victims, this algorithm provides law enforcement with enough data to begin a reasonable search, rather than expending a great deal of effort for little to no reward, or searching randomly based on data that has been improperly analyzed, and thus is essentially the same as a random search. Having this knowledge gives law enforcement officials a great deal of aid in preventing what often seems to amount to random crime.

**Appendix A**

**Map of Attacks by BTK Killer**

**Appendix B**



**Executive Summary**

When a serial criminal is targeting an area with violent crimes seemingly at random, law enforcement is naturally focused on attempting to find and stop this person before anyone else can be harmed. Our method for finding a serial criminal is both accurate and efficient. We compile two different schemes based on the knowledge of what the criminal has done in the past, in order to predict where they will strike next, as well as who may be the victim. Without a method such as ours, the search for a serial killer is nothing more than random guesses based on unorganized information and chance.

First we look at the location of previous attacks, and find the areas most dense with the work of the criminal. Since quite often criminals tend to center their crimes around a “home base,” the areas that are most dense are often close to the criminal's home or area of the next attack. With our method, analysts draw a circle around the two locations that are the furthest apart. Then they follow an algorithm that specifies how to draws smaller and smaller circles until the two closest events are connected. The area where the most circles intersect gives an indication as to where the criminal is most likely to strike next.

Next we take information about the demographics of the victims to attempt to determine whom the criminal is most likely to target. We then take advantage of census data to find the areas with the most people of this nature, relative to the rest of the population. This is one of the unique aspects of our method. We do not just look at one or two factors that the criminal may be targeting. Our equation takes all common traits into account, and weights them based on how prevalent they are among the victims. So, for example, if there is no distinct trend as to the gender of the victims, then these factors will not have a great weight in the larger analysis. However, if the criminal is targeting only women, then this becomes one of the most important factors.

After these values are calculated, we take a map of the city, or affected area, and overlay a map of the first scheme, with the areas where a crime is more likely to occur made darker than those with very little chance of a crime occurring. We then cover the map with a second overlay, where the areas with the highest ratio of the criminal's target demographic are the darkest. We then look at the combined overlays, and the area that is overall the darkest is where we predict that the criminal is most likely to strike next.

In order to use our method most effectively, the criminal must have already had five victims in order to collect enough data. Also, these victims should have some sort of demographic characteristic in common. If the criminal is truly targeting random victims, then our second scheme will have no darker areas, since every person is equally likely to be a target. Additionally, the criminal must be targeting a fairly specific geographic region.

Since this method must be efficient, it is recommended that members of law enforcement compile maps of their respective territories, as well as the relevant census data ahead of time. There is no point in using a method such as this, if by the time all of necessary data is obtained, the killer has already struck several times. Though this may require some substantial effort initially, once the information is compiled, the method becomes very simple and efficient.

One of the primary concerns with serial crime is that often it appears to be primarily random. Our method does not cover all of the possible scenarios that a serial criminal could create, and sometimes there may be more effective methods already in place, especially if more information is known. However, when nothing is known about the criminal, and the victims do have some traits in common, then our method is going to be the most efficient and most accurate option.