

# Measuring UK Crime Gangs

Giles Oatley  
Department of Computing  
Cardiff Metropolitan University  
Cardiff, UK  
Email: goatley@cardiffmet.ac.uk

Tom Crick  
Department of Computing  
Cardiff Metropolitan University  
Cardiff, UK  
Email: tcrick@cardiffmet.ac.uk

**Abstract**—This paper describes the output of a study to tackle the problem of gang-related crime in the UK; we present the intelligence and routinely gathered data available to a UK regional police force, and describe an initial social network analysis of gangs in the Greater Manchester area of the UK between 2000-2006.

By applying social network analysis techniques, we attempt to detect the birth of two new gangs based on local features (modularity, cliques) and global features (clustering coefficient). Thus for the future, identifying the changes in these can help us identify the possible birth of new gangs (sub-networks) in the social system.

Furthermore, we study the dynamics of these networks globally and locally, and have identified the global characteristics that tell us that they are not random graphs – they are small world graphs – implying that the formation of gangs is not a random event. However, we are not yet able to conclude anything significant about scale-free characteristics due to insufficient sample size.

## I. INTRODUCTION

There have been numerous studies of criminal networks and gangs; as highlighted in [1], the popularity of qualitative studies of gang-related issues soared during the 1980s and 1990s, following renewed media and public interest, statistical advances, and increased government funding. Qualitative studies have taken three major forms: (a) surveys of law enforcement officials (and at times other agency personnel) regarding gangs in their jurisdictions and actions taken to control them, (b) analyses of data compiled by law enforcement agencies and/or court officials, and (c) self-reports of samples of youth and/or young adults. There have been calls for research evidence to be drawn into police practice, but development of such an agenda has been hampered by a range of factors [2]. Research into youth gangs, especially the age at which youths join gangs and the early precursors, has been conducted in the USA and Canada [3], China [4] and Hong Kong [5].

However, the UK has been slow in carrying out research into gang crime (excepting work done in North East London [6] in 2007), and especially into what actions work best at controlling it. In Greater Manchester, a region in the north of the UK that has had a significant gun crime problem related to gang activity (primarily due to acute social deprivation in the area) [7]–[9], recent police initiatives have started to address this problem [10].

Social network analysis has been applied across a wide number of domains, providing a unifying language to describe

disparate systems ranging from social interactions to power grids. There is also a growing body of literature applied to crime analysis (for example: [11]–[17]). Related work [18], [19] on analysing the strength of weak ties in crime through steady state equilibria modelling has also been successful. Identifying structural holes, betweenness and social capital reinforces the value of using social network analysis for gang research [20].

We present the dynamics of a social network study of these gangs and their associates, using the intelligence gathered by police observations of known gang members and associated criminals. We develop the statistical analysis of network dynamics, combining well-known global topological measures, local motifs and modules [21]–[23]. Network motifs are sub-graphs that appear more frequently in a real network than could be statistically expected. At a global level, if these networks of associations exhibit clustering behaviour this indicates the presence of gangs. At a local level, any defined substructures will provide us information about the gang structure. We are interested in modelling the dynamics of the gangs, their development and fragmentation into new gangs, and we hope that the study of the dynamics in such modules will provide information on the structural changes within gangs that lead to birth of new gangs, and predictors of other gang-related behaviour.

Furthermore, we investigate if the networks have scale-free, small-world or other characteristics [23]–[25]; small-world networks are characterised by a diameter that grows logarithmically with their size. One important characteristic of the small-world phenomenon is that each pair of nodes are connected through a relatively small number of steps to a huge network size defined by the total number of nodes. Scale-free structures consists of many nodes with low degrees and a few hubs with high degrees [21], [22], [26]. If the offender networks can be classified into either (or both) of these categories (or other known network types), then this provides not only insight into the dynamics of the gang network, but also operational uses; for instance, network disruption/destruction strategies, nodes/offenders to monitor, and so on.

## II. PROBLEM DESCRIPTION AND DATA

Gun crime in Manchester first gained media attention in 1988 after concern over eight shootings and a gun-related murder, at a time when gun crime was considered rare in the UK. Nevertheless, gun crime in Manchester appears to have begun in the late 1970s at a time of rising unemployment and poverty in the area.

Numerous shootings – both fatal and non-fatal – have taken place over the years as the the Pepperhill, Gooch, Doddington and Longsight Crew gangs (see Table I) have clashed over drug territories and other disputes. Many of these gun fire exchanges were on public streets, some were planned acts and some were spontaneous events.

Gang label	Gang Name	Formation
A	Gooch	1990s
B	Doddington/Pepperhill	1990s
C	Longsight Crew	c.2001
D	Rusholme Crew Gangsters	c.2004

Table I. GANG NAMES AND APPROXIMATE DATES OF FORMATION.

In 2001, a new approach to tackling gun crime began to develop with police working more closely with the local community and other agencies. The Manchester Multi-Agency Gang Strategy (MMAGS), a multi-agency approach to tackling gun crime and deterring young people from entering into a gang/gun culture was initiated as a result of a UK Home Office report [27]. The report concludes that about 60 per cent of shootings are thought to gang-related, with violence in general, and gun violence and fatal shootings in particular are concentrated in specific small areas of South Manchester, and that gangs in South Manchester are loosely turf-based.

The geographical proximity of the gang locations and hub of all these activities can be seen in Figure 1, where the distance between Gangs A and B is hundreds of meters, literally a few streets away from each other. Gangs A and B show a negative attitude towards each other, often resulting in ‘tit-for-tat’ gun crimes. The alignment between Gangs A and D is possibly because of a mutual rivalry with B, while the positive alignment of B with C is because A has encroached on C’s ‘territory’ for drug sales. The gang locations are overlaid on the locations of all serious crimes (murder, attempted murder, manslaughter, kidnapping, serious wounding, and firearms offences) recorded in the data available to the consortium for the period 1980-2007. Agreeing strongly with the Home Office report [27] we find: 38% (n=162) of all serious crimes occurring within 1 km radius (of gang locations) and 63% of all serious crimes occur within 2 km, and 53% (n=9) of murders are within 3 km; 38% (n=34) of attempted murders are within 1 km and 63% within 2 km; and, 33% (n=17) of serious woundings are within 1 km and 48% are within 2 km.

### III. POLICE DATABASES

The database used for this analysis included the list of associates for each gang member, with fields such as unique identifiers for each offender, date of birth, relationship between the offenders, ethnic origin, reason reported and date of occurrence.

#### A. Link types

The network links available are quite different to other existing work with networks of burglars or retail fraudsters [14], [28]). Examples of the data (link types) from which the networks of offenders are developed can be found in Table II. These link types are: *Accomplice*; *Brother-Brother*; *Boyfriend*; *Brother*; *Sister*; *Charged with*; *Child*; *Cohabitant*; *Foster child*; *Foster parent*; *Friend*; *Girlfriend*; *Guardian*; *Other*; *Parent*;

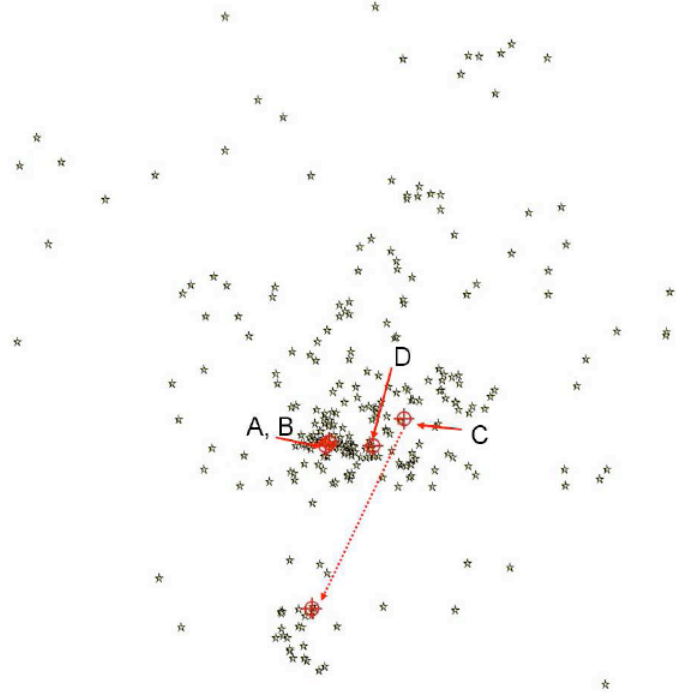


Figure 1. All serious crimes: murder, attempted murder, manslaughter, kidnapping, serious wounding, firearms offences. Gang C has moved into an additional location with drug selling. Gang geographical locations. *positive* indicates a positive alignment between the gangs, *negative* indicates negative alignment.

*Relative*; *Spouse*; *Sister-Sister*; *Ward*; *Gay Boyfriend*; and *Gay Girlfriend*.

An explanation of the dataset from Table II follows, and it is clear that it is a rich source of information. However, there are also many inconsistencies, and if this data is to be used to its full potential it will require a great deal of pre-processing, using natural language processing, matching with regular expressions, information extraction, and so on. There should also be further categorisation applied, considering that 50% of the data is classified of type *Other*.

#### B. Observations and inconsistencies in the dataset

The following indices refer to rows in Table II, for instance *1-i* refers to *1. Accomplice* from the *Relationship* column, and *(i.) Arrested Together* from the *Reason Reported Examples* column.

- *1-i* and *2-ii* indicate that the data is not rigidly recorded or categorised
- *1-iii* is incorrectly categorised
- *3-i*, *6-ii* and *9-i* illustrate that the intelligence is fallible, and is often based upon beliefs, and also that the link types are not all of equivalent strength, for instance the strength of a *Belief* link (possibly false) versus a *Charged Together* link (definitely true)
- *4-i* and *5-i* illustrate how the same information can be described, often in different forms, in separate fields

Magnet Category	Relationship	Frequency/Percentage	Reason Reported Examples
Crime related	1. Accomplice	502, 10.7%	(i.) Arrested Together (ii.) Believed To Be Dealing Drugs Together (iii.) X's Sister Is Y's Girlfriend
	2. Charged with	45, 1.0%	(i.) Charged Together Murder (ii.) Arrested Together
Familial	3. Brother	65, 1.4%	(i.) Believed To Be Half-Brothers
	4. Child	23, 0.5%	(i.) Father & Son (ii.) Admitted Above Named Is His Dad
	5. Parent	20, 0.4%	(i.) Mother & Son
	6. Relative	173, 3.7%	(i.) Cousins (ii.) X States Y Is His Uncle
	7. Sister	18, 0.4%	(i.) Brother And Sister (ii.) Stated They Are Brothers
	8. Spouse	2, 0.0%	(i.) Arrested Together Handling
Friendships	9. Cohabitant	5, 0.1%	(i.) Possibly Living Together At Anon Street
	10. Friend	1409, 30.0%	(i.) Stop Checked Together In car (ii.) Attended Club Together (iii.) Seen Together
	11. Girlfriend	61, 1.3%	(i.) Have Child Together
	12. Boyfriend	10, 0.2%	(i.) Girlfriend/Boyfriend
Other	13. Other	2364, 50.3%	(i.) Ex-Boyfriend Of The Above Named (ii.) R Claimed E Stabbed Him (iii.) C Intends Killing A/N Re Murder Of Bros (iv.) Tog At Nightclub, Oldham (v.) Seen Together (vi.) Attended Murder Trial (vii.) Arrested Together In Anon (viii.) D's Number In C's Mobile (ix.) Seen Together At Moss Side Festival

Table II. EXAMPLES OF THE 'ASSOCIATES' DATA. THIS DATA IS USED TO CREATE THE SOCIAL NETWORKS. GANG MEMBERSHIP COMPRISES A MIX OF SAME-AGE LOCAL FRIENDSHIP GROUPS, BLOOD RELATIVES AND RECRUITS: UK HOME OFFICE REPORT [27].

- 7-ii shows an obvious mistake with *Brothers* recorded in the *Sisters* category
- 8-i contains not only information about cohabitation, but also intelligence about handling stolen goods
- 11-i illustrates again that links can be stronger or weaker, for instance the child may mean that there is a stronger bond/link between the offenders
- 10-i-iii could all be placed in the *Other* category
- 13-ii,iii contain a lot of intelligence
- 13-v is a weak form of link, and should really indicate whether it was on good or bad terms
- 13-vii should be in either the *Accomplice* or *Charged With* categories
- 13-viii is noteworthy as it is a very specific link, a mobile phone link

### C. Limitations of the data and affects to SNA

Overall, it is concerning that this data is used as a gang database, but without explicit qualifications. Furthermore, it is generally not purged, but membership would not necessarily

have an effect on sentencing. Comparing to gang criteria by states in the USA, '*Identified by reliable source (police)*', and '*associates with members*' would secure membership in Florida [29].

Criticism of gang databases ranges from the position of being 'unconstitutional' if they are not correctly maintained, for instance, not regularly purged of citizens who have left the gang world [30], to including inaccuracies:

*"In sum, gang databases appear to be riddled with factual inaccuracies, administrative errors, lack of compliance with departmental guidelines, and lack of oversight. But this is not the worst of it. The root of the problem may be that even if properly applied, application of the subjective criteria would not produce useful results."* [31]

It is important to be critical of information about gangs that come from the police or from journalists, which is often based on impressions and not on thorough research. For instance 'intelligence' that describes that there are leaders in gangs who are responsible for 'recruitment' is at odds with our findings, that our network data does not find any obvious leadership (which is in line with many criminological studies on gangs). Various network outcomes contradict current stereotypes of

gang behavior, for example the existence of many links and intermediaries between different and sometimes conflicting gangs.

#### IV. IDENTIFYING COMMUNITY STRUCTURE

In order to investigate community structure we removed any nodes with less than six connections (i.e. degree 6); Figure 2 shows data from 2002, with the well-established Gangs A and B, and also the newly formed Gang C (in 2001). The Gangs A, B, and C are highly interconnected, with Figure 2 also showing the ‘go-between’, labelled as  $ab^*$  and  $bc^*$ . Individuals who are only connected to one gang, and who are highly connected within themselves, are labelled  $a^*$  and  $b^*$ . In this way it is easier to see the communities.

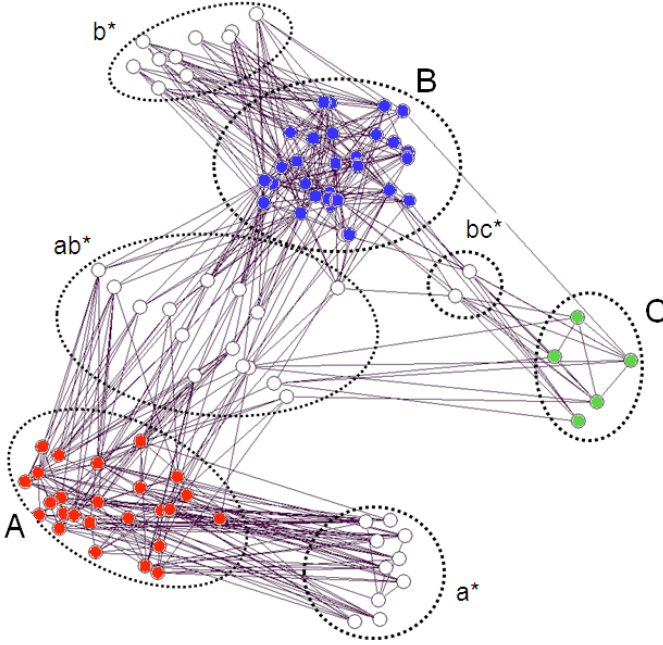


Figure 2. Link reduction, showing Gangs A and B and emergence of Gang C (for 2002). This also illustrates the large amount of non-gang members who are associated with individual gangs ( $a^*$ ,  $b^*$ ) or who are intermediaries ( $ab^*$ ,  $bc^*$ ).

#### V. NETWORK CHARACTERISATION

A series of experiments were carried out to determine how the gang networks compare with well-known networks, for example scale-free and small-world networks.

##### A. Small-world networks

Table III presents the clustering coefficient [32] (CC) for each individual year, alongside the node and edge counts and various other measures to describe the network. For any simple connected graph  $G$  with at least two vertices, the clustering coefficient (1-neighbourhood) [32] measures the extent to which vertices linked to any given vertex  $v$  are also linked to each other. Or in other words, are the friends of my friends also my friends? This is 1-neighbourhood clustering. The clustering coefficient 2-neighbourhood is a less stringent condition, and states: of the friends of my friends, are they linked to me by other friends?

The links presented in Table III are cumulative; that is, the links and nodes for 2002 include not only the new links and nodes for 2002, but also those for 2001 and 2000.

Measure	2000	2001	2002	2003	2004	2005	2006
Number of nodes (n)	1095	1295	1487	1752	2090	2229	2408
1/n	0.00091	0.00077	0.00067	0.00057	0.00048	0.00045	0.00042
4/n	0.00365	0.00309	0.00269	0.00228	0.00191	0.00180	0.00166
log(n)	6.999	7.166	7.305	7.469	7.645	7.709	7.787
log(log(n))	1.95	1.97	1.99	2.01	2.03	2.04	2.05
Number of links	1565	1903	2295	2844	3540	3872	4265
Total possible links	598965	837865	1104841	1533876	2183005	2483106	2898028
Diameter	12	14	11	11	14	12	13
Average path length	4.85	4.82	4.68	4.57	4.86	4.78	4.70
Density	0.00261	0.00227	0.00208	0.00185	0.00162	0.00156	0.00147
Betweenness	0.107	0.117	0.172	0.205	0.146	0.102	0.100
CC (cumulative)	0.47	0.48	0.47	0.46	0.49	0.55	0.56
CC (per year)	0.24	0.57	0.34	0.15	0.62	0.25	0.30

Table III. NETWORK MEASURES FOR 2000-2006. CLUSTERING COEFFICIENTS ARE ALWAYS GREATER THAN  $4/N$ . AVERAGE PATH LENGTHS ARE ALWAYS LESS THAN  $\log(N)$ .

Table IV shows the same network measures, but this time the data has been sliced into the members of the Gangs A, B, C and D.

Measure	A	B	C	D
Number of nodes (n)	859	617	431	513
1/n	0.00116	0.00162	0.00232	0.00195
4/n	0.00466	0.00648	0.00928	0.00780
log(n)	6.76	6.42	6.07	6.24
log(log(n))	1.91	1.86	1.80	1.83
Number of links	844	1047	602	707
Total possible links	368511	190036	92665	249571
Diameter	7	5	6	7
Average path length	3.61	3.38	3.37	4.11
Density	0.00396	0.00550	0.00648	0.00537
Closeness	0.302	0.298	0.393	0.305
Betweenness	0.185	0.179	0.350	0.239
CC	0.16	0.19	0.15	0.12

Table IV. NETWORK MEASURES FOR GANGS A, B, C, D. CC IS THE AVERAGE CLUSTERING COEFFICIENT FROM [32], CONSIDERING ONLY 1-NEIGHBOURHOOD.

A small-world network has both local connectivity and global reach [32], and is a simple connected graph  $G$  exhibiting two properties:

- 1) Small characteristic path length: the presence of short-cut connections between some vertices results in a small characteristic path length  $L(G)$ .
- 2) Large clustering coefficient: each vertex of  $G$  is linked to a relatively well-connected set of neighbouring vertices, resulting in a large value for the clustering coefficient  $C(G)$ .

To determine whether our network is a random one or is small-world, we can test whether or not it has exponential  $k$ -connectivity distribution. We do not observe this in the data, however, we do see large clustering coefficients, and the average path lengths are always less than  $\log(n)$ . Based upon these two criteria we can still conclude that our networks have small-world characteristics.

##### B. Scale-free networks

This section also refers to the preceding tables, where we find a mixture of evidence for and against the case for scale-free networks.



Plotting the clustering coefficient as a function of the number of nodes  $n$ , should follow the power-law distribution for scale-free networks (see later experiments), with the clustering coefficient being roughly four times larger than random networks [26]. The value of the clustering coefficient for a random networks will be  $1/n$ . In this way we are able to compare the values of  $4/n$  against CC in Tables III and IV. As the cumulative links increase from 2000 to 2006, the value of CC generally increases (with the number of nodes  $n$ ) and is always significantly higher than the values of  $4/n$ . Each of the gang values for CC are also significantly higher than would be expected in a random network.

The diameter of the network (longest path length) should be approximately  $\log(\log(n))$  for scale-free networks. In both cases (for the gangs and the years) the real values are significantly higher than would be expected for a scale-free network.

The average path length should be approximately  $\log(n)$  for scale-free networks. For both the ‘years’ and ‘gangs’ data it was actually smaller than  $\log(n)$ , indicating scale-free networks.

The statistics on degree centrality were low, indicating that there is no group leader. As we know when Gangs C and D are formed (2001 and 2004 respectively), it is interesting to note that the characteristic of the networks at this time are that the betweenness centralisation reaches 0.2. It is necessary to compare the closeness and betweenness averages for each gang against the value for the overall network.

### C. Power law investigation

This section examines whether the data would follow the power-law distribution for scale-free networks, and therefore we plotted the clustering coefficient as a function of the number of nodes  $n$ .

**Definition 1.** A quantity  $x$  obeys a power law if it is drawn from a probability distribution:

$$P(x) \propto x^{-\alpha}$$

where  $\alpha$  is a constant parameter of the distribution known as the exponent or scaling parameter. The scaling parameter typically lies in the range  $2 < \alpha < 3$ .

Our initial power law investigations used a log-log plot and  $R^2$  values, and these all produced  $\alpha$  values within this typical range (between 2 and 2.5). However being roughly straight on a log-log plot is a necessary but not sufficient condition for power-law behaviour [33], and that there are problems (bias and inaccuracy) with fitting to the power-law distribution using graphical methods based on linear fit on the log-log scale.

We therefore proceeded to use maximum likelihood estimation (MLE), which is a far more robust method for estimating the scaling exponent [33], [34]. We report the maximum likelihood estimate of the scaling exponent ( $\alpha$ ), the estimate of the lower bound of the power-law ( $x_{min}$ ).

By optimising the Kolmogorov-Smirnov goodness-of-fit statistic, we can use a *goodness of fit* to estimate where the empirically-best scaling region begins [33]. Given an observed data set and a hypothesised power-law distribution from which

the data are drawn, we can then test whether our hypothesis is a plausible one using the goodness-of-fit test (the Kolmogorov-Smirnov statistic), given the data, and generate a p-value that quantifies the plausibility of the hypothesis.

Employing the Kolmogorov-Smirnov test we are able to choose among the hypotheses that:

- $H_0$ : the data follow a specified distribution;
- $H_a$ : the data do not follow the specified distribution.

We did not use Vuong’s test to check for alternative distributions (non-power-law distributions) which could have produced the data. Instead, because our sample sizes are small (i.e.,  $< 100$ ), we explicitly used an experimental finite-size correction, as recommended by [33].

Figure 3 shows our results for our network between 2000–2006. In all cases the exponent  $\alpha$  is less than 2. Only when the power-law exponent is in the range  $2 - -3$  do the hubs tend to connect to form a single cohesive hierarchy [35]. The goodness-of-fit (gof) and p-values however are significant. Even though the p-values are above 0.1 (arbitrary threshold level), we err on the side of caution because of the low  $\alpha$  value and the small sample size. When  $n$  is small, meaning  $n \leq 100$ , we cannot rule out the power-law hypothesis [33]. It is possible, for small values of  $n$ , that the empirical distribution will follow a power law closely, and hence that the p-value will be large, even when the power law is the wrong model for the data [33].

Table V shows our results for the power law exponent for the different gangs against years. The case is similar in that there are significant gof and p-values, however in nearly all cases the exponent is less than 2, and again we did not test for alternate explanatory distributions, satisfied (operationally) that the the tail was heavy in all cases, indicating the presence of very well connected offenders.

Gang	1999	2000	2001	2002	2003	2004	2005	2006	2007
A	<b>2.65</b>	1.47	1.00	<b>1.91</b>	1.07	0.10	0.94	0.77	0.74
B	<b>2.95</b>	1.44	<b>3.64</b>	<b>1.88</b>	1.36	0.09	0.97	0.63	0.51
C	0.27	0.24	0.17	0.32	0.38	0.02	0.46	0.36	0.32
D	1.26	0.76	0.56	0.69	1.14	0.03	1.21	0.81	0.65

Table V. POWER LAW EXPONENTS FOR GANGS, AGAINST YEARS. SIGNIFICANT RESULTS ARE SHOWN IN BOLD-FACE.

Based on these experiments we are therefore unable to comment whether the networks possessing scale-free characteristics, however we can conclude that we have small-world networks, since consistently there are larger clustering coefficients and shorter path lengths compared to a random network with same number of gang members. This means two things for our system:

- The smaller path length means that the criminal activity (contagion) spreads more easily in this network than in a random network.
- Larger clustering coefficient means that contacts of contacts are treated as contacts as well.

### D. Emergence of gangs

We might see changes in the path length and clustering coefficients from 2000 to 2005, indications of how the gangs

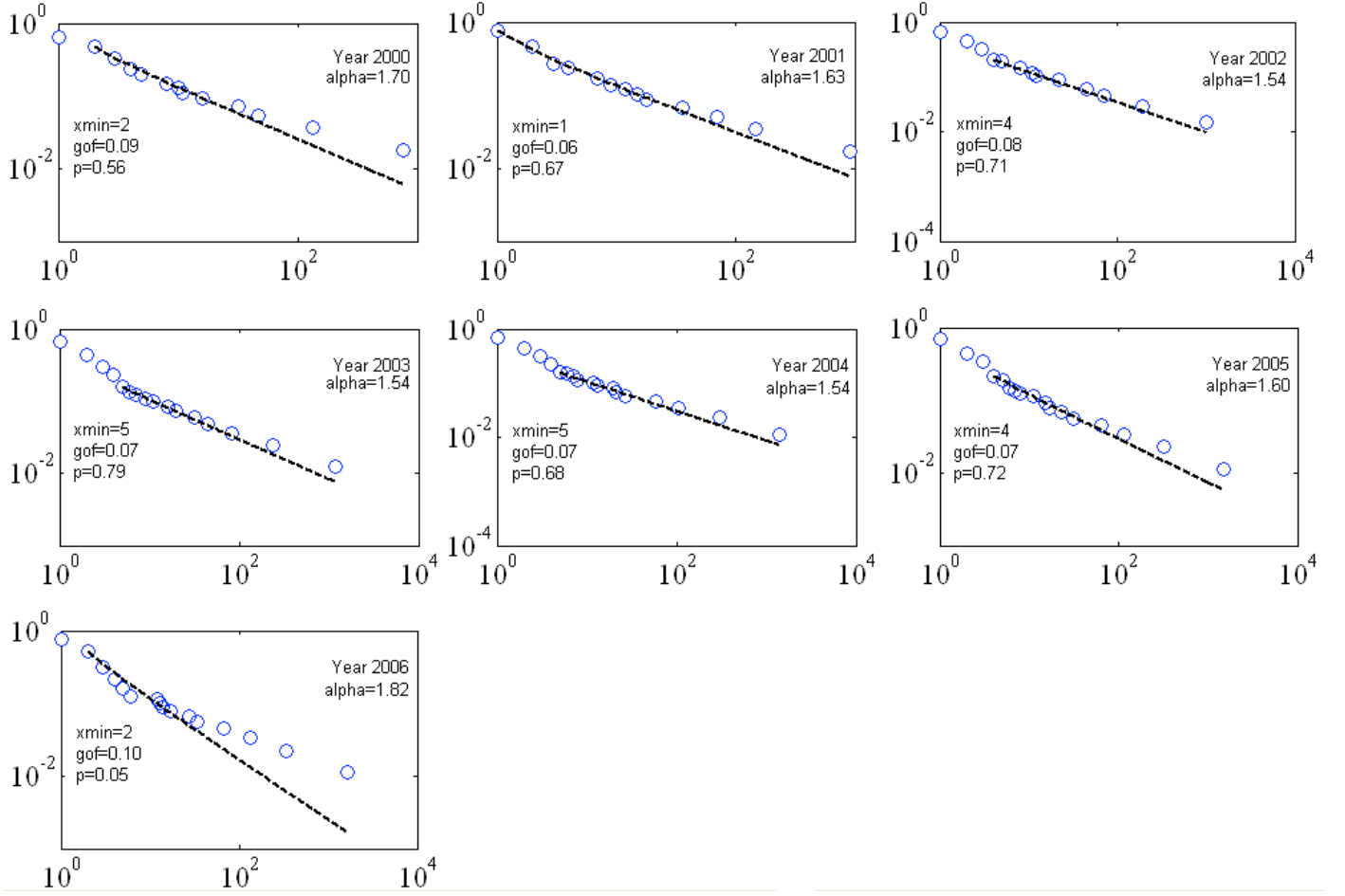


Figure 3. Power law investigations. A power law is fitted to each years data and various statistics calculated: the exponent  $\alpha$ , xmin, goodness-of-fit (gof) and p-value.

have become more closely knit or are splitting apart. By examining annual links for 2001 and 2004, we might predict that the cumulative links decrease and the annual links increase, just before/as a gang forms, then both values increase afterwards as everyone becomes linked together. This is not the case, and neither are we able to see any meaningful behaviour in these data.

Figures 4 and 5 show the clustering coefficients for each gang and against years, and is also a pictorial view of the new links per year. In Table 5 the CC value of each gang dips at 2004. What this may indicate is clustering due to non-gang members (from Figure 2, offenders who are connected to gang members:  $a^*$ ,  $b^*$ ,  $bc^*$  and  $ab^*$ ) and less clustering that previous years between members of gangs themselves. There is also a significant peak in clustering during 2001 for Gang B, whereas all other gangs suffer a decrease in clustering.

## VI. DISCUSSION

The model of two rival sets of gangs is potentially a misrepresentation of the much more complex sets of smaller cliques and fluid changes within the larger gang structures. However, the four gangs discussed do exist, and are the main gangs; what is not possible is a high degree of exactitude.

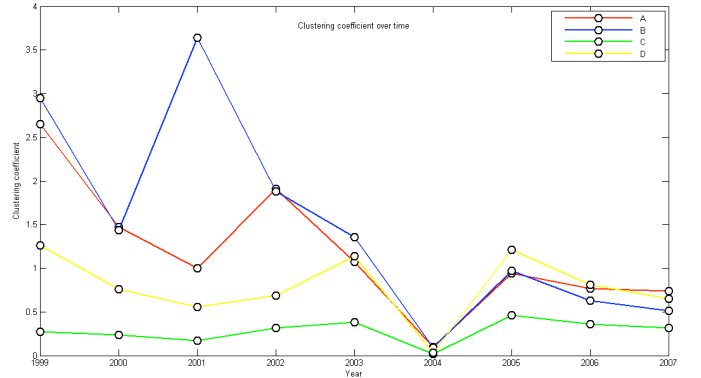


Figure 5. Per year clustering coefficients for each gang. Gang C was formed in 2001, Gang D in 2004.

### A. External and internal factors

It is difficult to determine through the gathered data what is happening in the networks. We have little recorded evidence of gang formation, even knowing when these events ‘allegedly’ occurred, similarly with the alleged ‘melt-down’ following the death of prominent gang leader Raymond Pitt in 1995 [36]. The links are based upon observations by police officers – do

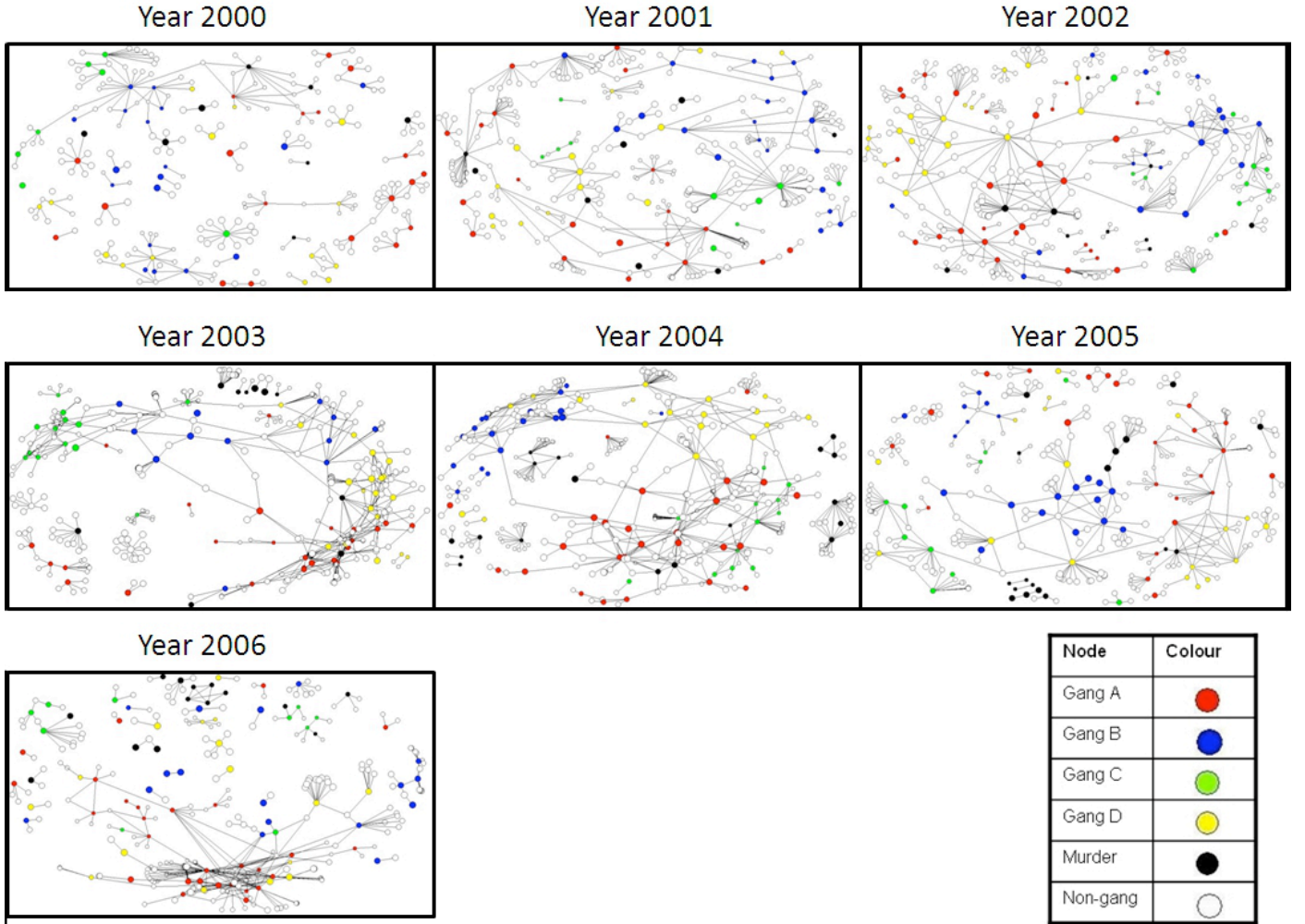


Figure 4. Annual links formation. Only nodes directly connected to a gang member are included. The network measures for each of these networks can be found in Table IV.

we expect that these complex social situations can be reflected in the reported links? Can we detect these events, and did they really happen as they have been passed down to us? We are in the difficult situation of using intelligence instead of concrete facts, and this intelligence is often a poor reflection of what is happening in the chaotic social world of gang culture.

We require a much better analysis of link types, developed a model where individuals learn about crime opportunities by interacting with other peers; for instance whether weak ties play an important role in explaining criminal activities [19], especially gang homicide [37].

The theoretical predictions of the model are confirmed by the empirical analysis since they find that weak ties, as measured by friends of friends, have a positive impact on criminal activities.

Furthermore, for 2001 and 2004, it would be interesting to examine the kinds of links within each gang which split apart.

#### B. Covert links

Data collection is very partial and certainly biased, since not every actor is exposed to an equal extent and therefore

some of those observed (perhaps the ‘usual suspects’) contribute far more to the dataset than others. Our earlier observation of a decrease in clustering as the network temporarily fragments, before an increase in clustering as everyone becomes linked together (as commented upon in Section V-D) finds equal explanation through the police having intelligence on the formation of a new gang and actively seeking observations on this event.

## VII. CONCLUSIONS

The work presented in this paper contains our initial findings about the offender/gang networks in Manchester in the UK, using network analysis. The uses of this technology in an operational context are significant. Even using the networks merely as visual representations of otherwise cognitively unmanageable data contained in spreadsheets and databases is operationally very useful, for knowledge sharing and training, and identifying key offenders. When further pre-processing is carried out, and the quality of the data collection process is improved, there will be significant future work available with this dataset.

The police crime recording database is routinely gathered

and available for analysis. The additional databases of histories and associates of gang offenders are routinely gathered by the UK's National Crime Agency<sup>1</sup>, who investigate gang and gun-related crimes. These data sources are potential rich sources of information for computer science technologies to deliver crime prevention and detection decision support systems.

Criminal behaviour (modus operandi and offence profiling) is to be incorporated into social network analysis. This approach uses retrospective methodologies, appropriate given the time scale and the pilot nature of most work. Future work must be given the resources to adopt prospective methods in order to increase the validity of decisions concerning their contribution.

#### ACKNOWLEDGMENTS

We acknowledge the UK's Engineering & Physical Sciences Research Council (EPSRC) Sandpit on Gun Crime (September 2005, Warwickshire, UK), funded by the IDEAS Factory; and the assistance of Xcalibre, the Greater Manchester Police's specialist gang crime task force.

#### REFERENCES

- [1] L. A. Hughes, "Studying Youth Gangs: Alternative Methods and Conclusions," *Journal of Contemporary Criminal Justice*, vol. 21, no. 2, pp. 98–119, 2005.
- [2] K. Bullock and N. Tilley, "Evidence-Based Policing and Crime Reduction," *Policing*, vol. 3, no. 4, pp. 381–387, 2009.
- [3] K. G. Hill, C. Lui, and J. D. Hawkins, "Early Precursors of Gang Membership: A Study of Seattle Youth," *Juvenile Justice Bulletin*, 2001.
- [4] V. J. Webb, L. Ren, J. Zhao, N. He, and I. H. Marshall, "A Comparative Study of Youth Gangs in China and the United States: Definition, Offending, and Victimization," *International Criminal Justice Review*, vol. 21, no. 3, pp. 225–242, 2011.
- [5] T. W. Lo, "Triadization of Youth Gangs in Hong Kong," *British Journal of Criminology*, vol. 52, no. 3, pp. 556–576, 2011.
- [6] J. Pitts, "Reluctant Gangsters: Youth Gangs in Waltham Forest," University of Bedfordshire, Tech. Rep., 2007.
- [7] BBC News, "Gangs and guns – who takes the rap?" [http://www.bbc.co.uk/manchester/have\\_your\\_say/2003/01/06/gun\\_crime.shtml](http://www.bbc.co.uk/manchester/have_your_say/2003/01/06/gun_crime.shtml), April 2003, (accessed 2014-04-01).
- [8] —, "Gun gang members 'die by age 24'," <http://news.bbc.co.uk/1/hi/england/manchester/3907049.stm>, July 2004, (accessed 2014-04-01).
- [9] G. Hales, C. Lewis, and D. Silverstone, "Gun Crime: The Market In and Use of Illegal Firearms," Home Office, Home Office Research Study 298, 2006.
- [10] BBC News, "Greater Manchester gun crime down by 20%," <http://www.bbc.co.uk/news/uk-england-manchester-11563463>, October 2010, (accessed 2014-04-01).
- [11] S. W. Baron and D. B. Tindall, "Network structure and delinquent attitudes within a juvenile gang," *Social Networks*, vol. 15(3), pp. 255–273, 1993.
- [12] P. Klerks, "The network paradigm applied to criminal organizations," *Connections*, vol. 24(3), pp. 53–65, 2001.
- [13] L. L. Hansen, "Girl 'crew' members doing gender, boy 'crew' members doing violence: An ethnographic and network analysis of Maria Hinojosa's New York gangs," *Western Criminology Review*, vol. 6(1), pp. 134–144, 2005.
- [14] G. Oatley, J. Zeleznikow, R. Leary, and B. Ewart, "From links to meaning: A burglary data case study," in *Proceedings of the 9th International Conference on Knowledge-Based Intelligent Information and Engineering Systems (KES 2005)*, ser. Lecture Notes in Computer Science, vol. 3684. Springer, 2005, pp. 813–822.
- [15] G. Oatley, B. Ewart, and J. Zeleznikow, "Decision support systems for police: Lessons from the application of data mining techniques to 'soft' forensic evidence," *Artificial Intelligence and Law*, vol. 14(1-2), pp. 35–100, 2006.
- [16] A. Calvó-Armengol and Y. Zenou, "Social networks and crime decisions: The role of social structure in facilitating delinquent behavior," *International Economic Review*, vol. 45(3), pp. 939–958, 2004.
- [17] C. E. Hutchins and M. Benham-Hutchins, "Hiding in plain sight: Criminal network analysis," *Computational & Mathematical Organization Theory*, vol. 16(1), pp. 89–111, 1995.
- [18] A. Calvó-Armengol, T. Verdier, and Y. Zenou, "Strong and weak ties in employment and crimes," *Journal of Public Economics*, vol. 91(1-2), pp. 203–233, 2007.
- [19] E. Patacchini and Y. Zenou, "The strength of weak ties in crime," *European Economic Review*, vol. 52(2), pp. 209–236, 2008.
- [20] A. V. Papachristos, *Studying Youth Gangs*. AltaMira Press, 2006, ch. Social Network Analysis and Gang Research: Theory and Methods.
- [21] L. F. Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas, "Characterization of complex networks: A survey of measurements," *Advances in Physics*, vol. 56(1), pp. 167–242, 2007.
- [22] M. O. Jackson, *Social and Economic Networks*. Princeton University Press, 2008.
- [23] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45(2), pp. 167–256, 2003.
- [24] D. J. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton University Press, 2003.
- [25] R. Albert and A. L. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74(1), pp. 47–91, 2002.
- [26] R. Albert, I. Albert, and G. L. Nakarado, "Structural vulnerability of the North American power grid," *Physical Review E*, vol. 69(2), p. 025103, 2004.
- [27] K. Bullock and N. Tilley, "Shootings, Gangs and Violent Incidents in Manchester: Developing a Crime Reduction Strategy," Home Office, Crime Reduction Research Series CRRS13, 2002.
- [28] G. Oatley, K. McGarry, and B. Ewart, "Offender network metrics," *WSEAS Transactions on Information Science & Applications*, vol. 12(3), pp. 2440–2448, 2006.
- [29] J. Barrows and C. R. Huff, "Gangs and Public Policy: Constructing and Deconstructing Gang Databases," *Criminology and Public Policy*, vol. 8, no. 4, pp. 675–703, 2009.
- [30] J. B. Jacobs, "Gang Databases: Context and Questions," *Criminology and Public Policy*, vol. 8, no. 4, pp. 705–709, 2009.
- [31] J. D. Wright, "The Constitutional Failure of Gang Databases," *Stanford Journal of Civil Rights & Civil Liberties*, 2005.
- [32] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393(6684), pp. 440–442, 1998.
- [33] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-law distributions in empirical data," *SIAM Review*, vol. 51(4), pp. 661–703, 2009.
- [34] M. L. Goldstein, S. A. Morris, and G. G. Yen, "Problems with fitting to the power-law distribution," *The European Physical Journal B – Condensed Matter and Complex Systems*, vol. 41(2), pp. 255–258, 2004.
- [35] L. A. Andamic, R. M. Lukose, and B. Huberman, *Handbook of Graphs and Networks: From the Genome to the Internet*. Wiley, 2003, ch. Local Search in Unstructured Networks, pp. 295–317.
- [36] P. Walsh, *Gang War: The Inside Story of the Manchester Gangs*. Milo Books, 2005.
- [37] A. V. Papachristos, "Murder by Structure: Dominance Relations and the Social Structure of Gang Homicide," *American Journal of Sociology*, vol. 115, no. 1, pp. 74–128, 2009.

<sup>1</sup><http://www.nationalcrimeagency.gov.uk/>