Predicting the Development of Juvenile Delinquency by Simulation

Tibor Bosse, Charlotte Gerritsen, and Michel C.A. Klein

Vrije Universiteit Amsterdam, Department of Artificial Intelligence
de Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
{tbosse, cg, mcaklein}@few.vu.nl

Abstract. A large number of delinquent activities are performed by adolescents and only occur during this period in their lives. One of the main factors that influence this behaviour is social interaction, mainly with peers. This paper contributes a computational model that predicts delinquent behaviour during adolescence based on interaction with friends and classmates. Based on the model, which was validated based on empirical data, the level of delinquency of pupils is simulated over time. Furthermore, simulation experiments are performed to investigate for hypothetical scenarios what is the impact of the division of students over classes on the (individual and collective) level of delinquency.

Keywords: social simulation, social learning, delinquent behaviour.

1 Introduction

One of the main challenges in Criminology is to understand, explain and predict when individuals show delinquent behaviour [4]. Obviously, there is a wide range of potential contributors to the emergence of crime, varying from environmental opportunities to social influences. In this paper we focus on the latter. Learning (delinquent) behaviour by social interaction is something that is often observed in adolescents [11]. During the period from 12 to 18 year old, people are more susceptible to the opinion of their peers. In some situations, their desire to be part of a group can be so strong that they break some rules to achieve this desire. This is consistent with the theory by Moffitt [10] who states that one can divide delinquents roughly into two groups, namely life-course persistent offenders and adolescence limited offenders. The behaviour of the first group is caused by neuropsychological problems during childhood that interact cumulatively with their criminogenic environments across development, which leads to a pathological personality. This behaviour will usually continue through life. Instead, the behaviour of adolescence-limited offenders is caused by a gap between biological maturity and social maturity. It is mainly caused by mimicking antisocial role models like peers, but also parents and school are important contributors. These offenders peak sharply at about age 17 and drop fast in young adulthood.

In this paper we exploit simulation techniques to study the development of such juvenile delinquency. As mentioned above, this type of behaviour is limited to a certain period of time, and some of its direct causes are clearly determined. This provides opportunities to develop a computational model of this process. In previous research [2], we developed such a model, which was able to predict the level of delinquency of students based on information about the personal characteristics and their peer network. The model was validated by using a large dataset with information about 1730 scholars (taken from [14]).

The main contribution of the current paper is to show how this model can be used to perform so called *what-if simulation experiments*. In these simulations the existing (validated) model is applied to a hypothetical situation, which is slightly different from the situation in the existing empirical data. For example, we want to see what happens to the level of delinquency (both of individuals and of the classes) when the composition of the classes is altered. Interesting questions here are, among others:

- What is the effect when we put the most delinquent students together in one class?
- Is it better to spread the delinquent and non-delinquent students equally over classes?

To answer such questions, this paper proposes to make use of social simulation techniques [3]. In recent years, a number of papers have successfully tackled criminological questions using social simulation, e.g., [7, 9]. However, the current paper differs from these approaches in that we do not attempt to reproduce existing data, but rather explore how hypothetical scenarios would evolve. We will create these hypothetical scenarios by making small modifications in existing scenarios (e.g., change the composition of classes), and run the simulation model on the modified data. The main question that we would like to answer is whether the composition of a school class has an influence on the overall level of delinquency of the pupils. This is an interesting topic, since it is often believed that the structure of schools and peer networks has an important impact on juvenile delinquency [8, 12].

The paper is organised as follows. In Section 2 we describe how the data used for the simulation experiments were collected. The simulation model itself is presented in Section 3, and the experiments in Section 4. Finally, Section 5 concludes the paper with a discussion and some ideas for future work.

2 Data Collection

The model presented in this paper is based on empirical data from a longitudinal research project. This research was performed by the Netherlands Institute for the Study of Crime and Law Enforcement (NSCR) in the so called 'School Project' [14], which focused on peer network formation, personal development, and school interventions in the development of problem behaviour and delinquency.

In this project, a large number of high school students were surveyed by means of questionnaires. As respondents, a cohort of students was used that started high school during the school year 2001/2002. The first year of secondary education in the Netherlands is comparable with 7th grade in the United States (most students are 12 or

13 years old). These students were surveyed during three consecutive years: 2002, 2003 and 2004.

During these three years, the respondents had to fill out a number of questionnaires. Their delinquent behaviour was measured using self-reports of a variety of offences. The self report method is a standard procedure in Criminology, and it results in fairly reliable estimates of delinquency levels of young people, when it is conducted in a proper way and in an anonymous setting. Respondents were asked if they had ever committed an offence and, if so, how often during the reference period. The measures of self-reported delinquency used in this study come from 12 questions, among which: in the last year, how many times did you: "paint graffiti", "vandalise property", or "steal small things from shops worth less than 5 Euros" The total delinquency measure indicates how many types of the 12 possible types of delinquent behaviours were reported by the person.

The respondents also had to answer a number of questions about their friends (e.g. with whom they spent a lot of time, who were their best friends), to obtain information about their social networks. In the analyses, friends' numbers were linked to the respondent's own number, enabling the networks of friends to be mapped and analysed.

Further, the study also used a substantial number of other measures on risk factors that are central in criminological theories and have been found to correlate with delinquency in the past (e.g. low supervision and support by parents, low bond with school, low law conformity, high impulsivity, high temper). For more details of the empirical research see [13, 14].

3 Simulation Model

In this section the simulation model used for the experiments is described. First, in Section 3.1, the methodology behind the design of the model is discussed (based on [1]. Section 3.2 presents the implementation of the model, and Section 3.3 shows how the original model was extended in order to incorporate information about classes.

3.1 Design Methodology

As a first step in the process of designing the model, an initial dynamic model was developed for the development of delinquency through social learning in a class room, based on an analysis of the literature (see Figure 1). A more detailed description of this model is provided in [1]. The model describes the influences of several personal characteristics, as well as the influences of other peers. More specifically, the delinquency of an agent is influenced by its previous delinquency, its individual personality traits (e.g. temper, impulsiveness), and external factors (i.e., the school, the parents, and peers). This original model has the form of a set of differential equations, where delinquency is measured as a real number between 0 and 1. In [1], it has been shown that this model can be used to simulate delinquency development of a small set of agents in a classroom. The simulations exhibited several patterns that would be expected based on the criminological literature.

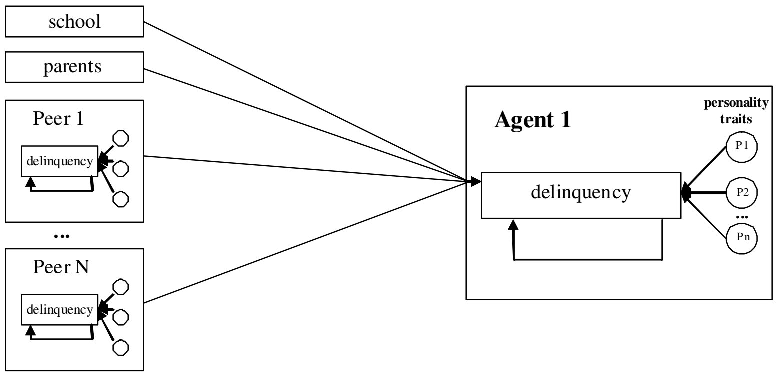

Fig. 1. Overview of the simulation architecture (from [1])

A next step was to validate the model based on the empirical data mentioned in Section 2. In this research [2], a representative sample of the collected dataset has been selected, and has been split up in a training set and test set. Each set contained the data of around 250 pupils. A lot of pupils were left out of the original dataset, because their questionnaires were not suitable. This was caused, for instance, by gaps in the answers or because they were only attending a particular school during part of the research period. When making this split, we guaranteed that there was no overlap between the schools used in the training set and those used in the test set. We developed an evaluation method that could be used to quantify the correctness of models and to discriminate between accurate and less accurate models. This measure accommodates the intuitive ideas about a correct prediction in one number (see Section 3.2). The model was calibrated with the data in the training set by taking the model from [1] extended with some additional factors reported in [14], and systematically adjusting it and comparing the simulation results with the actual measurements in the training set (scaled to a number between 0 and 1). The adjustment consisted of both ignoring factors in the model (i.e. leaving out variables in the formulae) and calibrating parameters (i.e. changing the value of weighting variables), thereby creating different variations of the model.

Finally, the second data set was used to validate the different variations of the model that seemed promising during the calibration phase. In this phase, we did not change the model or parameters, but just calculated the accuracy according the developed measure for all formulae that resulted in a high score in the first phase. This method gives an unbiased validation of the accuracy, as the validation is performed on a different data set than the tuning.

3.2 Implementation

To implement the model, we used standard numerical simulation software. A ‘school class’ was modelled as a multi-dimensional array, where each array represented a different student. The different dimensions represented characteristics of the students over time. For example, these dimensions specified the individual characteristics (like impulsivity and risk-orientedness) and the relations to peers. To calculate the new delinquency of each agent, the following algorithm was used (in pseudo code):

```
For each agent:
1. determine current delinquency
2. determine individual characteristics
3. compose the social network (friends)
4. calculate average delinquency of social network
5. calculate new delinquency, using information from step 1, 2, and 4
```

To calculate the new delinquency of the individual agents (step 5), various variants of the model have been tried, each incorporating some of the factors identified in the previous section. These different models are depicted in Table 1. For example, model variant 1 (a baseline model), always predicts that students will not become delinquent. The last column denotes the accuracy rate for each model, which was calculated as follows:

Accuracy Rate = (w*Hits + Correct Rejections) / (w*Hits + w*Misses + False Alarms + Correct Rejections)

where Hits, Misses, Correct Rejections and False Alarms are defined according to the classical measures in signal detection theory [5]. For more details, see [2]. The factor 'risk orientedness' (model 8 and 9) indicates the extent to which the pupils like performing exciting activities, and the factor 'deviance reinforcement' (model 9-11) indicates the extent to which the pupils are sensitive to influences of their friends.

As can be seen, variants 10 and 11 have the highest accuracy. This means that the previous delinquency combined with the impulsivity, the level of deviance reinforcement by friends, and the delinquency of (best) friends, seem to be the best predictors for delinquent behaviour.

Table 1. Variants of the model and accuracy values (taken from [2])

Model variant	Main factors used	Accuracy
1	always predict non-delinquency	45.42
2	always predict delinquency	54.58
3	randomly predict non-delinquency and delinquency (with ratio 1:1)	50.31
4	randomly predict non-delinquency and delinquency (with ratio 3:1)	48.16
5	all factors identified in Section 3.1	66.10
6	delinquency last year	70.52
7	delinquency last year, delinquency friends	71.04
8	delinquency last year, delinquency friends, risk-orientedness, temper	72.64
9	delinquency last year, risk-orientedness, deviance reinforcement, delinquency friends	73.62
10	delinquency last year, impulsivity, deviance reinforcement, delinquency best friends	76.41
11	delinquency last year, impulsivity, deviance reinforcement, delinquency friends, delinquency best friends	76.24

In addition to the accuracy, the quality of the models has also been tested using a Relative Operating Characteristics (ROC) analysis. This method has been used because this is a standard measure in the literature and allows us to compare the results with studies in other domains. The outcome of this analysis is a curve which represents a graphical plot of the fraction of true versus the fraction of false positives for a binary classifier system as its *discrimination threshold* is varied (see Figure 1). The threshold in our model is the value of the calculated delinquency above which a pupil is classified as delinquent. We calculated the area under the ROC curve (AUC), a scalar measure for the quality of the predictions, for each model. For model variant 10, the AUC is 0.79. An AUC-value larger than 0.70 is called ‘acceptable’, larger then 0.80 ‘excellent’ and larger then 0.90 ‘outstanding’ [6][1].

3.3 Incorporating Class Information

Although model 10 and 11 produce the highest accuracy rates, these model variants are not particularly appropriate for the aims of the current paper. That is, the goal of this paper is to predict for hypothetical scenarios (which are slightly different from the existing situation) how the delinquency of the students would have developed. And since it is not very realistic to assume that one can easily modify, say, the impulsivity or the friend network of students, variant 10 and 11 are not very useful candidates for these ‘what-if experiments’.

For this reason, two additional variants of the model have been developed. These models (variant 12 and 13) use the composition of classes. For obvious reasons, in practice it is much easier to manipulate students’ class composition than their friend networks. Therefore this factor was also manipulated within the hypothetical scenarios. To this end, variants of the model have been developed that take the delinquency of class members into account.

Model variant 12 predicts that a student will become delinquent if (s)he was delinquent in the previous year OR (s)he is part of a delinquent class AND (s)he has a high value for ‘deviance reinforcement’. Here, being part of a delinquent class is defined as the situation that the average delinquency of all students in the class is higher than a certain threshold. Note that this variant does not make use of the friend network.

The ROC curve obtained for this model variant 12 is depicted in Figure 2a, when compared with a random model (variant 3). As can be seen, variant 12 performs much better than the random model. The AUC of model variant 12 was 0.734, and its accuracy is 72.33. Although this is lower than the AUC and accuracy of variant 10 (resp. 0.79 and 76.41), we decided to use variant 12 for the simulation experiments described in the next section, because (as explained above) this variant contains the students’ classes as one of the factors.

In addition, a model variant has been developed that also takes the delinquency of the friends into account. Variant 13 predicts that a student will become delinquent if (s)he was delinquent in the previous year OR delinquency of the friends times the ‘deviance reinforcement’ is higher than a certain threshold OR (s)he is part of a delinquent class AND (s)he has a high value for ‘deviance reinforcement’. For being

[1] Note that the AUC approach both has advantages and drawbacks when compared to the accuracy approach. An advantage is that this measure is rather common in the literature, which makes it easier to interpret the numbers, and to compare them with other models. A drawback is that the resulting numbers are calculated on the basis of all possible instances of the discrimination threshold, whereas in the accuracy approach only the best instance is taken. And since for the simulation experiments only this best instance will be used, the accuracy approach could be considered to be more useful.

part of a delinquent class the same definition is used as in variant 12. The AUC of this model variant (see Figure 2b) is 67.52[2], and its accuracy is 72.66.

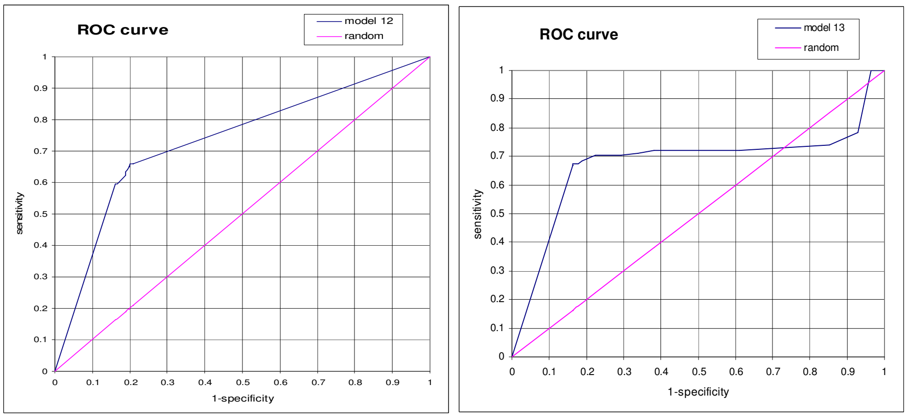

Fig. 2. ROC curves for a) model 12 and b) model 13 against a random prediction

4 Simulation Experiments

This section describes the simulation experiments that were performed to investigate the development of the delinquency of the pupils in the dataset for hypothetical scenarios. In Section 4.1, the setup of the experiments is explained. The results of the experiments are discussed in Section 4.2.

4.1 Approach

In the simulations, we compared the results of the simulation of the delinquency over one year using the actual class composition with the results of two simulations using a hypothetical composition, namely 1) a scenario in which all delinquent pupils are put together in the same class, and 2) a scenario in which all delinquent pupils are evenly distributed over all classes in a school. The goal of the comparison is to find out whether the change in the delinquency of pupils is positively or negatively influenced by the class composition.

The simulations are performed using three schools in our dataset, consisting of 6, 8 and 4 classes, respectively. These schools were chosen because many pupils of these schools filled out the questionnaire, so much data was available. In total 194 pupils were involved in the simulations. The simulations for the actual class composition and

[2] This relatively low number is mainly due to the dip at the right-hand side of the graph. This dip is caused by the fact that for some extreme values (which will obviously not be used in the simulation experiments) of the discrimination threshold, the model scores very bad. For this reason, in this case the accuracy may be more informative (see also footnote 1).

the two hypothetical scenarios have been performed two times, using each variant of the model that takes the class information into account (variant 12 and 13).

4.2 Simulation Results

Table 2 gives an overview of the development of the delinquency over a year according to the simulation with model variant 12 and 13. The first two columns indicate, respectively, the code of the school class in the study (e.g., '1 - 2' stands for 'class 2 of school 1'), and the amount of pupils in the class. In the next 3 columns, the 'base before' column shows the number of delinquent pupils in the actual class composition at the start, and the columns 'base after v12' / 'v13' the predicted number of delinquent pupils after a year using model variant 12 or 13 respectively. Similarly, the 6 subsequent columns show the number of delinquent pupils in a class at the start and the end using the hypothetical class compositions (called scenario 1 and 2). It can be seen that in scenario 1 all delinquent pupils of a school are put together in a class, while in scenario 2 the delinquent pupils are more or less evenly distributed over the classes.

Table 2. Results of the simulations of delinquency of pupils with alternative class compositions using model variant 12 and 13.

school class	class size	base before	base after v12	base after v13	scen1 before	scen1 after v12	scen1 after v13	scen2 before	scen2 after v12	scen2 after v13
1 - 1	7	0	0	0	6	6	6	1	1	1
1 - 2	17	2	2	2	0	0	0	1	1	2
1 - 3	10	2	2	2	0	0	1	1	1	1
1 - 4	6	1	1	1	0	0	0	1	1	1
1 - 5	6	1	1	1	0	0	1	1	1	1
1 - 6	6	0	0	0	0	0	0	1	1	1
2 - 1	9	2	3	3	9	9	9	5	6	6
2 - 2	17	7	9	9	17	17	17	5	7	7
2 - 3	11	6	6	8	9	10	10	5	5	7
2 - 4	11	6	6	7	0	0	1	5	5	6
2 - 5	10	3	3	5	0	0	3	5	5	7
2 - 6	1	0	0	0	0	0	0	0	0	0
2 - 7	17	9	9	10	0	0	4	5	5	6
2 - 8	13	2	2	2	0	0	1	5	6	6
3 - 1	6	1	1	1	6	6	6	3	3	3
3 - 2	10	2	2	2	7	7	7	3	3	3
3 - 3	21	7	7	7	0	0	0	4	4	4
3 - 4	16	3	4	4	0	0	1	3	3	4
total	**194**	**54**	**58**	**64**	**54**	**55**	**67**	**54**	**58**	**66**

As can be seen in Table 2, the difference between the baseline and the different scenarios is not very high. For model variant 12, the total number of delinquent pupils increases in scenario 1 from 54 to 55 instead of to 58 for the baseline, and in scenario 2 it increases as much as in the baseline. In model variant 13, the number of delinquent pupils increases to 64 in the baseline, while it increases to 67 in scenario 1 and to 66 in scenario 2.

In the simulations using variant 12 we see that the increase of the number of delinquent pupils is less for the scenario in which all bad guys are put together (scenario 1) than in the baseline scenario or the scenario in which the delinquent pupils are evenly distributed. However, this pattern is not visible when using model variant 13. Overall, the differences between the results of the baseline scenario and the two other scenarios are very small. Although care should be taken not to draw too strict conclusions from these preliminary experiments, this may be an indication that the use of alternative class compositions has little effect.

5 Conclusion

In this paper, we have presented a number of simulation experiments on juvenile delinquency. The simulations were performed using an existing model that was based on the theory of social learning. In our previous research we have used empirical data about juvenile delinquency and social networks to develop and validate this simulation model. In the current paper we have presented some novel variants of this model. Moreover, we have used the model to investigate the effect of different class compositions on the development of the delinquency in the total group of pupils.

The experiments show no significant difference between the change in the total number of delinquent pupils in the different scenarios. The two different scenarios represented two extreme situations: all delinquent pupils put together, or all delinquent pupils distributed over all classes. Therefore, our tentative conclusion is that the composition of classes has not so much effect on the overall development of the delinquency of the pupils in a school. This is an interesting finding, since it is often argued that careful composition of school classes is very important to prevent development of juvenile delinquency [8, 12].

However, there are a few remarks that can be made about our experiments, which could be of influence on this conclusion. First of all, the model is possibly not very precise (see the relatively limited accuracy) because of small size of the training set. It could be the case that with a more precise model (derived from a larger training set) stronger effects would be visible. A second remark concerns the size of the classes. The ones used in the simulated scenarios are much smaller than regular classes; as a consequence, the influence of other pupils in the class is smaller in our simulations than in reality. The class size is this small because the data of many of the pupils was not suitable, e.g., because of missing information or because they switched between schools. The fact that we do not see a clear effect could also be caused by the fact that the number of offenders in our data set is relatively small. Therefore, also the number of predicted changes will always be quite small. Finally, we want to remark that the current models do not allow pupils to learn non-delinquency from their peers at school, they can become delinquent. Although this apparently follows from our dataset in the best predictive models, it might be the case that this is a bit different in reality.

Despite these remarks, the approach presented in this paper has proved to be a useful additional tool for criminology scientists, as also confirmed by our colleagues in the Criminology department. The approach allows for experiments that can not be easily performed in the real world and could give some indication of the expected effects of class compositions on juvenile delinquency.

Acknowledgement

The authors are very grateful to Frank Weerman of the Netherlands Institute for the Study of Crime and Law Enforcement for his willingness to provide the empirical data from the 'School Project', and for a number of fruitful discussions.

References

1. Bosse, T., Gerritsen, C. and Klein, M.C.A. (2009). *Agent-Based Simulation of Social Learning in Criminology*. In: Proc. of the Int. Conf. on Agents and AI, ICAART'09. INSTICC Press, 2009, pp. 5-13.
2. Bosse, T., Gerritsen, C., Klein, M.C.A., and Weerman, F.M. (2009). *Development and Validation of an Agent-Based Simulation Model of Juvenile Delinquency*. In: Proc. of the Int. Symposium on Social Intelligence and Networking, SIN'09. IEEE Computer Society Press, 2009, pp. 200-207.
3. Davidsson, P. (2002). Agent Based Social Simulation: A Computer Science View. *Journal of Artificial Societies and Social Simulation*, 5(1).
4. Gottfredson, M. and Hirschi, T. (1990). A General Theory of Crime. Stanford University Press.
5. Green, D.M., Swets J.A. (1966) *Signal Detection Theory and Psychophysics*. NY: Wiley.
6. Hosmer, D., S. Lemeshow (2000). Applied logistic Regression. NY, John Wiley & Sons.
7. Liu, L., Wang, X., Eck, J., and Liang, J. (2005). Simulating Crime Events and Crime Patterns in RA/CA Model. In F. Wang (ed.), *Geographic Information Systems and Crime Analysis*. Singapore: Idea Group, pp. 197-213.
8. Matsueda, R.L. & Anderson, K., The dynamics of delinquent peers and delinquent behaviour. Criminology, 36 (1998) pp. 269-308.
9. Melo, A., Belchior, M., and Furtado, V. (2005). Analyzing Police Patrol Routes by Simulating the Physical Reorganisation of Agents. In: Sichman, J.S., and Antunes, L. (eds.), *Multi-Agent-Based Simulation VI, Proc. of the 6th Int. Workshop on Multi-Agent-Based Simulation, MABS'05*. LNAI, vol. 3891, Springer Verlag, 2006, pp 99-114.
10. Moffitt, T.E. (1993). Adolescence-Limited and Life-Course-Persistent Antisocial Behavior: A Developmental Taxonomy. *Psych. Review*, vol. 100, no. 4, pp. 674-701.
11. Sutherland, E.H., and Cressey, D.R. (1966). *Principles of Criminology*, 7th edition. Philadelphia: J.B. Lippincott.
12. Warr, M. Companions in Crime. The social aspects of criminal conduct. Cambridge, Cambridge University Press 2002.
13. Weerman, F.M., and Bijleveld, C.C.J.H. (2007). Birds of Different Feathers. *European Journal of Criminology*, vol. 4, issue 4, pp. 357-383.
14. Weerman, F.M., Smeenk, W., and Harland, P. (eds.), i.c.w. Ezinga, M., Slotboom, A.-M., Bijleveld, C., Laan, P. van der, and Westenberg, M. (2007). *Problem behavior of students during secondary education: Individual development, student networks and reactions from school* (in Dutch). Amsterdam: Aksant.