Analyzing Gang Reduction Strategies Through Dynamic Mode Decomposition

Yifan Li[a], **Jiazhong Mei**[b], **Aviva Prins**[a, 1], **and Qinyi Zeng**[a]

[a]University of California, Los Angeles; [b]University of California, Berkeley

This manuscript was compiled on September 29, 2018

We are interested in studying the effectiveness of the Gang Reduction and Youth Development (GRYD)'s prevention program by analyzing the change in estimated risk factor at different time steps. We view the change as a dynamical system of question responses per participant, and utilize the Dynamic Mode Decomposition algorithm to find the inherent temporal patterns. We are able to study youths' resiliency to change by their question responses, which may identify common themes that need to be addressed by the program. Furthermore, we develop a predictive model for the change in a youth's attitudes and behaviors every six months of participation in the program. This model yields a similar predictive power as a shallow neural network but with faster run time and increased interpretability. We also observe transient growth in the system, indicating the program may be effective for youth who initially increase in risky behavior.

dynamical system | gang reduction | transient growth | DMD

Introduction

The Los Angeles County's Office of Public Safety has made tremendous efforts to address gang violence in a comprehensive and coordinated manner though their Gang Reduction and Youth Development (GRYD) division. It is important to measure the effectiveness of these programs to aid social workers, youth, and the broader community.

GRYD has a prevention program that targets youth between the ages of 9 and 15 to dissuade them from joining gangs. Students considered to be low risk for joining gangs participate in the primary prevention program, which focuses on community development through after school activities. Students who exhibit risky behavior or attitudes are placed in the secondary prevention program. The secondary prevention program addresses individual needs through case management, family services, behavior modification therapy, and other services.

Eligibility for the program is determined by a questionnaire (Youth Services Eligibility Tool, or YSET) (1). The questionnaire contains behavioral and attitudinal questions in nine categories, such as parental supervision, impulsive risk taking, and peer delinquency. Scores from these categories are combined to calculate risk factor (RF), which ranges from zero to nine. Participants in the primary and secondary programs take the questionnaire every six months. Our dataset contains responses from all participants who have been in the GRYD program for at least one year between 2014 and 2017. Each person has at least three sets of responses with an interval of six months between them.

To investigate how the youths' responses change over time as a direct result of participation in the program, we model the dataset as a dynamical system. We utilize the Dynamic Mode Decomposition algorithm and provide results about the effectiveness of the program. We investigate the questions that change the least over the dataset and the youth who initially increase in negative behavioral and attitudinal characteristics.

Methods

Dynamic Mode Decomposition (DMD) is a method that studies the change in a dynamical system in terms of its growth or decay and spectrum. It was originally developed for analyzing fluid dynamics (2–6). The theory and algorithm are reviewed in Supporting Information, Section 2 & 3A. Controlled experiments are done on synthetic systems generated from GRYD data to verify the power of DMD on social science data (Supporting Information, Section 4). We get a list of eigenvalues λ and their corresponding eigenvectors ϕ from DMD. These are useful to identify transitions and periodic patterns in the system. Specifically, the real component $\Re(\log(\lambda))$, shows the growth or decay while the imaginary component $\Im(\log(\lambda))$, shows frequency and periodic patterns. The dominant eigenvalue and eigenvector pairs are selected by the computation of $\lambda|\phi|$ (4, 7), which is discussed in full detail in Supporting Information, Section 3C.

Results

We start by looking at the change in the distribution of risk factors among participants at different stages, shown in Figure 1a. After a year of participation in the program, youth tend to decrease in risk factor. The number of participants with low risk factor increases while the number of youth with high risk factor decreases. Both youth with initially low risk factor and students with high risk factor tend to decrease in risk. Youth with initial risk factor zero to three participate in the primary program, while youth with higher risk factor participate in the secondary program. There are significantly fewer youth with initial risk factor of three than four, which could indicate

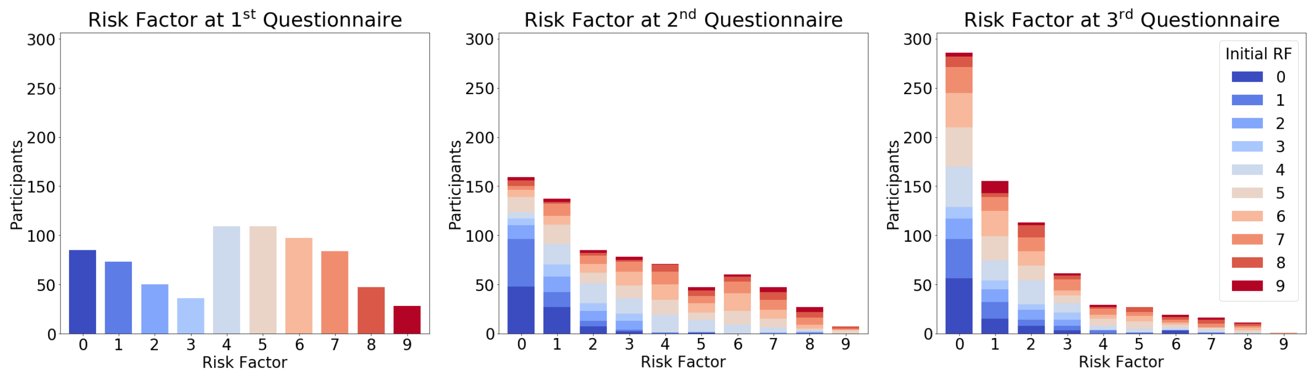
Significance Statement

We utilize an algorithm to identify the effectiveness of a gang prevention program that targets youth between the ages of 9 and 15. Our dataset contains anonymized demographic information and questionnaire responses on risky behavioral and attitudinal perspectives in six month intervals. We find that that the program is effective for each gender, ethnicity, and age group. We find certain attitudes are more resistant to change. We conclude that the program is effective, even for youth who increase in risky behavior.

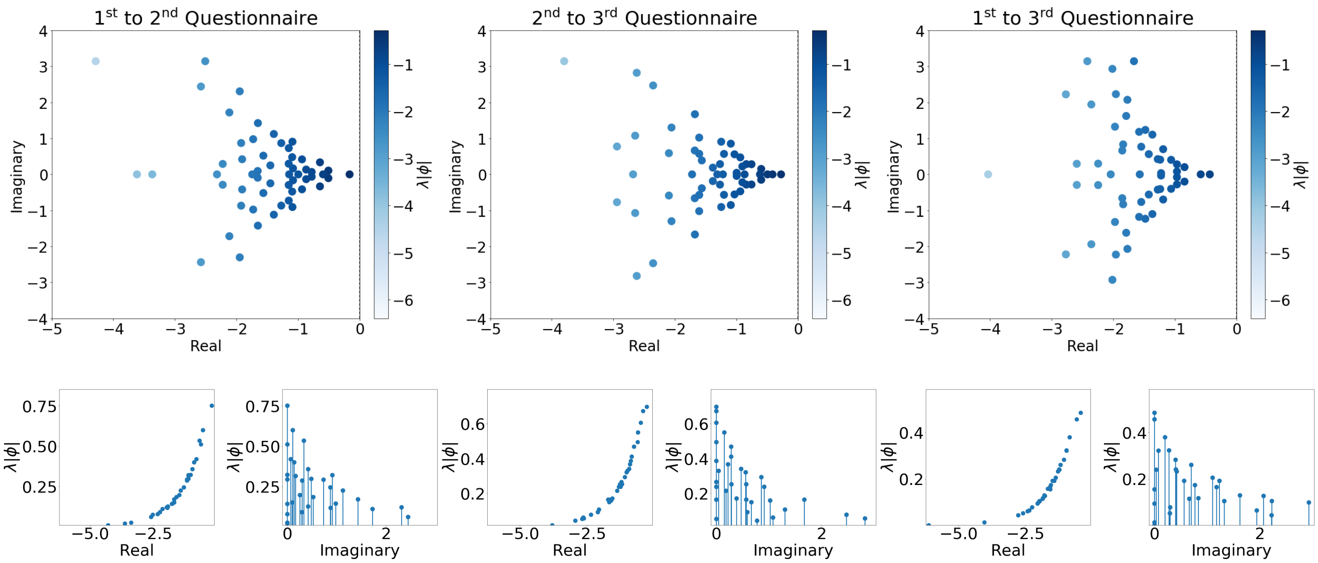
All authors contributed equally to this work.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: aviva@avivaprins.com



(a) Change in risk factor (RF) in the first year of participation in the prevention program ($n = 713$). Risk factor of each participant at intake, after six months, and after one year. The color corresponds to the participant's risk factor at the intake questionnaire.



(b) DMD on the first six months of participation. **(c)** DMD on the second six months of participation. **(d)** DMD on the first year of participation.

Fig. 1. Results from Dynamic Mode Decomposition, which returns eigenvalues λ and their corresponding eigenvectors ϕ. The real component, $\Re(\log(\lambda))$, shows the growth or decay while the imaginary component, $\Im(\log(\lambda))$, shows frequency and periodic patterns. Dominant features are determined by $\lambda|\phi|$. In the top plot, the x-axis is the growth, the y-axis is the decay, and color indicates significance. In the bottom plots, the x-axis is the growth or frequency (respectively) and the y-axis is the significance.

that the primary program struggles to retain its youth for a full year. Further analysis on the number of students allotted to each subprogram is required. Visually, the decay (i.e. the decrease in risk factor) is more pronounced in the second set of six months than in the initial six months.

Earth Mover's Distance of Risk Factors. To quantify the decay in the risk factors, we calculate the Earth Mover's Distance of the histograms shown in Figure 1 in order to calculate a decay score (as defined in Supporting Information, Section 1) for different periods of the program.

The decay score measures how much the participants' risk factors decay relative to the maximum level of decay possible. Theoretically, in a system that is not growing, the decay score ranges from 0 to 100, with 0 indicating no decay in risk factors and 100 indicating that the system experience extreme decay where all the risk factors become 0 at the end of the period. The computation tells us the participants who participate for two consecutive periods in the program (approximately one year) have a decay score of 62.6, which indicates a significant amount of decay of risk factors occurred from their first to their third take of the questionnaire. If we look at the two periods individually, during the first period in the program, the participants have a decay score of 33.8, which is considerably lower than that during the second period of program, which is 43.5. This suggests to us that perhaps the program is more helpful during the second period of participation compared to during the first period, and this could be caused by the lack of trust during the beginning stage of the intervention. Thus we expect to see decay in the system after we apply DMD to our dataset, and more specifically we expect to see that the decay shown from the DMD technique also captures the intensity of the decay during two different time periods. In the following sections, we will show that our results actually meet our expectations.

Decay and Frequency. We now apply DMD analysis discussed in Supporting Information, Section 3B on the GRYD data and study the change between evaluations in order to see if we observe a similar type of decay from the question responses. We first focus on the transition between first and second evaluations. Figure 1b shows the distribution of the indicator values exhibiting the growth or decay of the system after applying DMD. All these values lie on the negative side of the

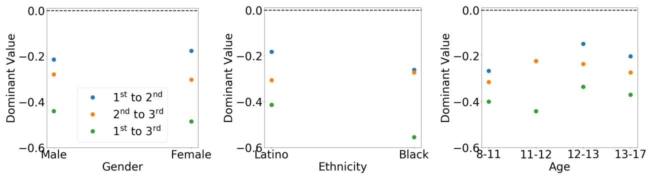


Fig. 2. Change in the dynamical system according to demographic information, using DMD with rank 10. Left: Gender (male, $n = 418$; female, $n = 282$). Middle: Ethnicity (Latino, $n = 537$; Black, $n = 135$). Right: Age (8-11, $n = 104$; 11-12, $n = 116$; 12-13, $n = 146$; 13-17 $n = 232$).

real axis, suggesting that there is decay in first six months of the program. In particular, the largest value is -0.165. The absence of an imaginary component of this value suggests that there is no apparent periodic pattern. Since the largest value stands out against other values, the overall decay in the system is dominated by this largest value.

The same analysis is then applied to the transition between the second and third evaluations, which corresponds to the effectiveness of GRYD program during the second six month period. From the scatter plot of eigenvalues in Figure 1c, we again see decay in the system, given that all values are on the negative side of the real axis. Notice that the largest value here is -0.276, which is smaller than that in the first six months. This suggests that there is more decay in the second six months than in the first six months, and that the duration of participation in the GRYD program is important. Furthermore, there is no apparent frequency in the system, which is consistent with the first system. Again, we identify that the overall decay in the system is dominated by the largest value.

Additionally, we perform our analysis on each gender, ethnicity, and age group to ensure GRYD is adequately addressing the youths' needs. As shown in Figure 2, both male and female participants decrease in risk factor in equivalent amounts. However, it appears that the Black youth are more responsive to the program than the Latino students, and older age groups are more resistant to the prevention program relative to younger age groups.

Attitudinal Resiliency to Change. In the previous section we have identified the largest eigenvalue and eigenvector pair as the dominant one in both systems. Here, we focus on the corresponding eigenvectors and interpret the contribution of each question in terms of its weight in the eigenvector. The weights in the eigenvector suggest how much the question contributes to the decay rate associated with its eigenvalue. The larger weight a question has in the dominant eigenvector whose eigenvalue is stable or has a small decay rate, the more resilient it is to decay.

Figure 3 shows the eigenvectors corresponding to the largest eigenvalue in each system. In particular, questions a2, c1, c4, c5, f5, f6, ij4 are identified from the transition between the first and second take of the questionnaire. As can be seen from the top plot, questions f5 and f6 have the largest weight among all questions, and they ask the participants "It is okay to beat people up if they hit me first." and "It is okay to beat people up if I do it to stand up for myself.". We see that there is an underlying theme associated with these questions, i.e. attitudes towards violence. Similarly, question a2 "I get very angry and 'lose my temper'." and c4 "Did you have a big fight or problem with a friend?" seem to be associated with violence. Questions c1 and ij4, however, are associated with the student's behavior at school.

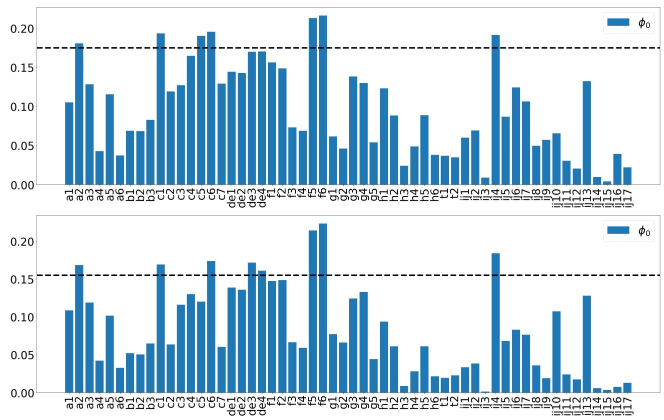


Fig. 3. Plots of the 56 YSET questions by their weight in the dominant eigenvectors. The top plot comes from transition between the first and second take of the questionnaire, the bottom plot from transition between the second and third take of the questionnaire.

The same theme and topics can also be identified from the eigenvector computed for the transition between the second and third evaluations. Questions a2, c1, f5, f6, and ij4 again have high weights. But more interestingly, questions de2 "I sometimes find it exciting to do things that might get me in trouble." and de3 "I often do things without stopping to think if I will get in trouble." have gained more weight compared to the first transition period. These two questions highlight students' negligence of the consequences from bad or even unlawful actions. This suggests that students during the second period of treatment are still reluctant to change their attitudes towards getting in trouble.

To conclude, both eigenvectors pick out a similar set of questions that have higher weights compared to others. We are able to identify some common underlying topics shared by these high-weighted questions, such as violence, school performance, and peer relationship. This suggests that although the GRYD prevention program is able to reduce participants' overall risk in joining gang, it is struggling to reduce the risk of youths in these topics mentioned above. The GRYD prevention program may wish to reevaluate the way they address the behaviors and attitudes we have identified here.

Prediction. Given an initial set of questionnaire responses, we can use Dynamic Mode Decomposition to predict how the questionnaire responses will change over time. To evaluate the accuracy of this prediction, we use 80% of our dataset to train the algorithm and use the remaining 20% to test its accuracy. In addition, we compare our results to a shallow neural network with one hidden layer. The error of these two models are shown in Figure 4. A shallow neural network is the most robust supervised machine learning technique applicable to this dataset due to its size.

As shown in Figure 4, both the training and testing dataset have a root mean squared error of approximately 0.3, regardless of the method. This indicates that DMD is comparable in accuracy to the most sophisticated machine learning techniques. DMD has the advantage of being more efficient. It takes half an hour to run the DMD algorithm for 100 iterations.

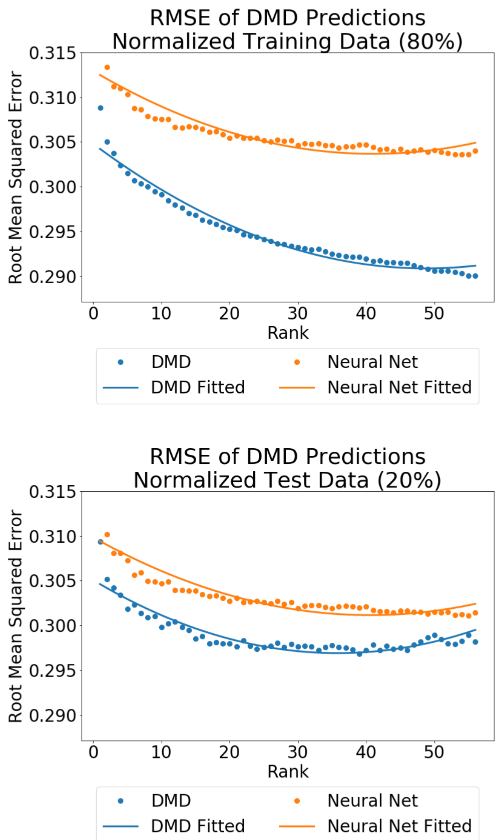


Fig. 4. Results from predicting 2,644 youths' progress after six months of participation in GRYD, using only their initial questionnaire responses. Dynamic Mode Decomposition and the neural network model are trained using 80% of the data. Top, the error of predicting this dataset. Bottom, the error of predicting a new dataset that was not used in the training process. The x-axis describes the range in possible parameters for the two models. DMD can take a rank parameter, which ranges from 1-56. The shallow neural net has 1-56 nodes in the hidden layer. The y-axis represents the root mean squared error of the two methods. For comparison, the RSME of a random walk on the initial questionnaire is 0.53.

To find the equivalent results with the neural network, seven hours are required, without the aid of a graphics processor. Finally, DMD provides other analytic information as discussed earlier in the paper while machine learning techniques do not.

Transient Growth. We also look at the dynamics in the system as a whole by tracking how an individual's response change through time. This can be represented by the energy of the system at different time step, as shown in Figure 5. Energy refers to the maximum possible value of response at each iteration. We see that the energy first spikes and then decreases as what we would expect. This transient growth in the system suggests that potentially there is some participants who will receive a higher score in their questionnaire responses after first six months before they start a decrease.

Discussion

The application of Dynamic Mode Decomposition to the GRYD dataset yields many informative results. We find that there is decay in the system, and a lack of periodicity, which is expected. We identify question responses that are resistant to change on all demographics and analyze the program's effectiveness on each demographic. Additionally, we are able to predict how the youth will modify in attitude and behavior after just six months in the program. This prediction is useful in understanding how the program affects different types of youth.

The dataset also includes youth who took the intake questionnaire multiple times but were ineligible for the program. In the future, we hope to compare these youth to the program participants. This will give us a better understanding of the dynamical system.

The question responses that are resilient to change also warrant further investigation. It may be that there is a common theme in these questions. Identifying the theme will help GRYD modify their program to address the needs of the program participants. In the future we hope to compare these questions with the previous work on topic modeling.

Additionally, we note the transient growth in the dataset. This indicates that even though a youths' risk factor may increase in the short term, the program is still effective. It may be that this is an artifact of the analysis, since risk factor does not completely describe the youth's attitudes. It is more likely that this is a feature of social science programs, since human behavior is complex.

Gang Reduction and Youth Development has an effective prevention program. From our analysis, the prevention program reduces youth's inclination to join gangs, provided that they participate in a full year of the program.

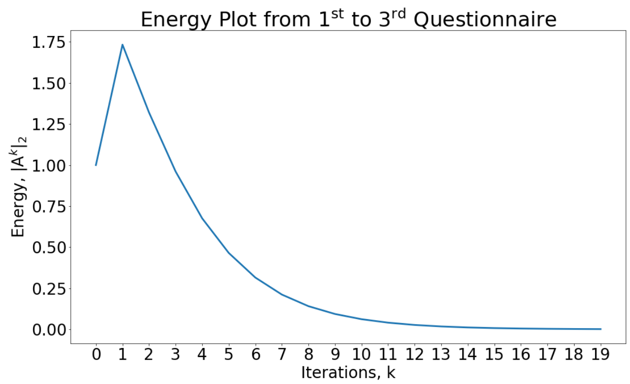


Fig. 5. Change in energy at each iteration. The energy increases at the first iteration before it decreases.

ACKNOWLEDGMENTS. We would like to thank Professor Andrea Bertozzi, Professor Jeff Brantingham, Dr. Omri Azencot, and Dr. Mike Lindstrom. Additionally, we thank Dr. David Arnold and Melanie Sonsteng for their help.

1. Hennigan KM, Kolnick KA, Vindel F, Maxson CL (2015) Targeting youth at risk for gang involvement: Validation of a gang risk assessment to support individualized secondary prevention. *Children and youth services review* 56:86–96.
2. Rowley CW, Mezić I, Bagheri S, Schlatter P, Henningson DS (2009) Spectral analysis of nonlinear flows. *Journal of Fluid Mechanics* 641:115–127.
3. Kutz JN, Brunton SL, Brunton BW, Proctor JL (2016) *Dynamic mode decomposition: data-driven modeling of complex systems.* (SIAM) Vol. 149.
4. Tu JH, Rowley CW, Luchtenburg DM, Brunton SL, Kutz JN (2013) On dynamic mode decomposition: theory and applications. *arXiv preprint arXiv:1312.0041.*
5. Schmid PJ (2010) Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics* 656:5–28.
6. Vallacher RR, Nowak A, Froehlich M, Rockloff M (2002) The dynamics of self-evaluation. *Personality and Social Psychology Review* 6(4):370–379.
7. Proctor JL, Eckhoff PA (2015) Discovering dynamic patterns from infectious disease data using dynamic mode decomposition. *International Health* 7(2):139–145.