Societal Impacts of Predictive Policing Algorithms in Crime Prevention

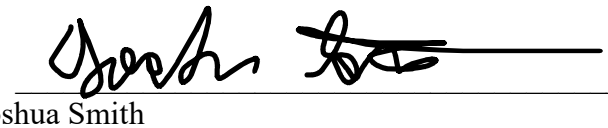
A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Joshua Smith
Spring, 2022

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature  Date 5/3/2022
Joshua Smith

Approved _____ Date _____
Hannah Rogers, Department of Engineering and Society

Abstract

As artificial intelligence and machine learning continue to develop and become applied towards our systems and technologies, we must remain cognizant of the impacts these new technologies bring. Namely, this paper analyzes the application of artificial intelligence and machine learning algorithms towards one of societies most integral systems, policing, known as predictive policing algorithms. The current and potential implementations of predictive policing algorithms were surveyed through academic research, and a determination was be made if predictive policing algorithms will become a positive benefit to the groups involved in its construction. Through the Social Construction of Technology (SCOT) framework, the paper argues that the development of predictive policing algorithm technology will stabilize in a manner that benefits all groups involved in its development.

Societal Impacts of Predictive Policing Algorithms in Crime Prevention

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science, School of Engineering

Joshua Smith
Spring 2022

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Hannah S. Rogers, Department of Engineering and Society

Introduction

Artificial intelligence (AI) and machine learning (ML) algorithms have redefined the world, fundamentally reshaping how modern systems and technology behave and "transforming every walk of life" (Allen & West, 2018). However, with the overwhelming potential AI and ML provide, it is even more important that issues of data privacy, security and bias are taken into account to ensure society is positively impacted by its application (Chui, et al., 2018). This research paper seeks to analyze the application of AI/ML algorithms towards predictive policing, known as predictive policing algorithms (PPAs), which aim to predict future locations and victims of crime using previously reported incident data for use by policing forces in crime prevention. PPAs are stated to greatly improve policing effectiveness, allowing for "policing that is smarter, more effective, and more proactive" and "allows police to make better use of limited resources" (Perry, et al., 2013, p. 1). However, this also comes with controversy over their implementation, with critics claiming that PPAs reinforce the current racial and economic bias present in policing. According to Amnesty International (2020), PPAs would reproduce or amplify any human biases that are inherent to the dataset used, which in the context of crime disproportionately affects certain races and socioeconomic status. Similarly, Selbst (2017) notes that software engineers may inadvertently design algorithms which discriminates against certain groups of people. Despite the criticisms facing PPAs, however, academics such as Yen and Hung (2021) note that, if their issues are addressed, PPAs are still highly effective in crime prevention and should continue to be used.

This research paper will seek to analyze interpretations of PPAs using the Social Construction of Technology (SCOT) framework developed by Bjiker, Hughes, and Pinch (1984). Interpretations of PPAs will be evaluated by exploring discourse around current PPA

implementations, their benefits, criticisms, and counterarguments to these criticisms. This will then culminate in a final SCOT analysis of these interpretations to determine if PPAs as a technology can stabilize in a manner that provides benefits to all social groups involved in its application.

Current Predictive Policing Algorithms

While there are a variety of different approaches and potential implementations of predictive policing algorithms, they can generally all be categorized into four distinct categories, as laid out by Perry, et al. (2013): methods for predicting places of crimes; methods for predicting individuals likely to commit crimes; methods for predicting offenders identities; and methods for predicting victims of crimes (Degeling & Berendt, 2017). Generally, most PPA implementations within the west fall within the first two categories, predicting crime locations and individuals likely to commit crime, and this paper will only focus on those two methods. While it is important to be aware that they exist, predicting victims and offenders of crime is much less feasible than the other two methods and are rarely used in regards to PPAs

When it comes to PPAs, the most widely-used implementation would be the location-based prediction of future crimes. Location-based predictions of crimes, as put forth by a location-based prediction software, PredPol (2021), are algorithms that work by using historical event datasets to train themselves and are then used to identify the times and places at highest risk of future crimes. Usually, these algorithms only make predictions based on past crime information given to them by the police, such as crime types, locations, and date/time, and nothing more. “No personally identifiable information is ever used. No demographic, ethnic or socio-economic information is ever used. This eliminates the possibility for privacy or civil rights violations seen with other intelligence-led or predictive policing models” (PredPol, 2021).

The police then use these location-based predictions to inform and optimize their patrol routes to ensure they are as effective and efficient as possible in preventing crimes (Degeling & Berendt, 2017).

The other widely-used PPA implementation, although much more controversial, is the approach of predicting individuals who are likely to commit crimes. These types of PPAs work by collecting various data points on an individual, including criminal history, education, employment, finances, etc., and then cross-referencing this data with previous offenders and non-offenders to determine how likely one is to become an offender (Perry, et al., 2013). For example, the Chicago 'heat list' compiles a list of names of people that are likely to commit crimes through analysis of arrested criminal networks, and police officers notify these individuals of any consequences should they engage in crime (Degeling & Berendt, 2017). However, a clear issue with predicting offenders is a question of ethics and civil rights, namely, the discrimination and abuses of power that can be borne from this approach; this will be explored further into the paper.

Benefit Analysis of Current Predictive Policing Algorithm Implementations

When it comes to the effectiveness of employing these PPA techniques in policing practices, there is seemingly strong statistical evidence towards PPAs being effective in crime prevention. For location-based predictive policing, research done by Mohler, et al. (2015), which performed two randomized controlled trials of real-time epidemic-type after sequence (ETAS) crime forecasting, found that ETAS models predicted 1.4-2.2 times more crime relative to a typical crime analyst with the same amount of data, and that predictions from ETAS models allowed patrols to reduce the volume of crime by an average of 7.4% relative to the norm. In other various research trials done into PredPol, research done by Bachner (2013) determined that

PredPol application in various trials over the years 2010-2011 resulted in a 27% drop in burglary incidents within designated patrol areas. Additional research done by Ferguson (2017 & 2020) saw PredPol accurately predict the location of 50% of gun homicides, and that it was able to perform better than typical human crime analysts with an overarching crime prediction rate of 4.7% versus the normal 2.1% of typical crime analysts within the same areas.

Overall, the research profile for location-based PPAs showcases a reduction in overall crime numbers due to their use. However, there are several caveats to consider with these results. While overall crime numbers are reduced, this may be due to police overly targeting the high-risk areas identified by the PPAs. This creates a situation where individuals within a certain area are disproportionately targeted, to either their benefit or detriment, and all other individuals outside of this area receive less police attention and possibly an increase in crime. A decrease in the aggregate crime numbers benefits the police, but it does not necessarily mean all individuals under their protection share the same, or equal, benefit. Still, these results still showcase that PPAs can enable the police to be much more effective in crime prevention over the aggregate.

The results of person-based PPAs, however, are much more nuanced. In fact, studies have shown person-based PPAs use result in no benefit at all. For example, research done by Saunders, et al. (2016) found that Chicago's 'heat list', also known as the Strategic Subjects List (SSL), identified individuals that were not more or less likely to be a victim or perpetrator of a homicide or shooting than anyone else; they were only relatively more likely to be arrested for a shooting, which was probably due to police officers targeting those on the list more. Not only does this mean the SSL provided no benefit to any group, it demonstrates that it enabled unfairly targeting certain individuals for arrest on no basis other than being on the SSL, which was already shown to identify nothing meaningful. Additionally, individuals on the SSL would

become part of the ‘Custom Notification Program’, which sends police and social services to selected individual’s homes, where they would receive a warning from police and program offers from social services (Rieke, et al., 2014). At best, this is offering social services and programs to those that need it, but this should already be determined by other technologies; at worst, this once again unfairly targets individuals for police scrutiny and surveillance with no basis. By and large, person-based predictive policing provides no measurable benefit to any group, the only possible exception being the profiting company of the algorithm, and only serves to negatively impact all groups, especially those it unfairly targets.

Criticisms and Counterarguments

Although we have shown how predictive policing algorithms (PPAs) have been implemented and that some have a demonstrable positive impact on crime prevention, many are critical of PPAs due to the potential societal ramifications they bring, especially during a time of increasing criticism towards policing practices. The first criticism, which can also be considered the lightest, is brought forth by Selbst (2017), noting that the software engineers who design PPAs may inadvertently design or implement their algorithms in a way which discriminates against certain categorizations of people. This criticism is amplified by Yen and Hung (2021), who note “it is difficult to detect the harm and find its cause with PPAs. It is thus difficult to assign responsibility for any harm caused by the deployment of PPAs.” Because of the difficulty to assign accountability with PPAs, versus other engineering disciplines, there is not much binding on the software engineer to ensure best intentions outside of their own ethical code and/or shareholders. However, there are easy solutions to this. Designing an algorithm without bias is simply an unobtainable reality, and is a problem not just with PPAs, but machine learning algorithms in general. As long as we design algorithms while keeping the scope in which the

algorithm has been empirically shown to be effective, and "choose a relatively good bias in both machine learning and societal norms" (Yen & Hung, 2021), then we can ensure algorithms will remain fair. And when it comes to the software engineers, engineering ethic guidelines are already present in all disciplines, and should be further expanded in enforced to ensure accountability in regards to PPAs, so that software engineers ultimately look out for the affected persons of their algorithms as opposed to delivering the best statistical results or profits to shareholders.

A second criticism towards PPAs is their potential and demonstrated discrimination towards certain groups of people due to either dataset bias or ineffective implementation. Academics have criticized the potential bias in the training data used by PPAs, which was already shown to be the case in Lum and Johndrow's research (2016), and that they directly contribute to racial and socioeconomic discrimination performed by police (Yen & Hung, 2021). For example, PredPol has stated that because no demographic, ethnic or socio-economic information is ever used by their algorithm, there is no possibility for privacy or civil rights violations as their absence would remove categorical discrimination (PredPol, 2021). However, according to Lum & Johndrow (2016), although it is perceived that algorithms produce unbiased results due to the 'neutrality' of computers and the exclusion of any variables which may introduce bias (race, income, etc.), the employment of algorithms actually retain and even amplify bias because training data is generated by a process that is inherently biased.

Based on Lum & Johndrow's research, they demonstrate that the most common approach of companies like PredPol to create 'race-neutral models', the exclusion of using race as a variable, still results in a racially biased outcome. However, in the same research done by Lum & Johndrow (2016), they point out you can solve the issue of racial and socioeconomic bias by

minimizing their existence via their anti-bias framework within data sets and algorithms without majorly impacting PPA effectiveness. While this means PPAs are currently racially and economically discriminatory, it is not an inherent part of PPAs, but simply an issue of their design, meaning future development of PPAs can rid themselves of this bias while giving the same benefits.

A third criticism academics point out are with the lack of transparency, awareness and peer-reviewed empirical research when it comes to PPAs. As put forth in the paper by Shapiro (2017), current predictive policing implementations are very closeted when it comes to their approach, and this causes much concern considering the amount of variance that is possible when it comes to what types of crime the algorithms can predict or block out, and the amount of bias that is possible if the algorithm or data is not suitable. In a similar research paper, Meijer and Wessels (2019) conclude that it is impossible to come to reasonable conclusions in regard to current implementations of PPAs due to the fact that there is a severe lack of empirical research that's been done into the field of predictive policing, and much of its supporting evidence is entirely anecdotal. While current implementations only have rare exceptions to this criticism, such as PredPol showing and explaining their implemented algorithm on their public website, this is also another criticism that is not inherent to PPAs and can be redeveloped to better benefit society. For example, PPAs could be made much more transparent and brought in-line with the public agenda, by having public hearings, increasing government oversight into PPAs, and funding much more research into PPA effectiveness and potential downsides.

Social Construction of Technology Framework Analysis

As has been noted in regards to the discussion of PPA, it is still an early and growing technology and thus the most important thing to consider is its development and continued use

moving forward. In order to understand this, the social construction of PPAs will be investigated using the SCOT framework analysis. Within this model, which is visualized in Figure 2, each group provides its own resources and feedback of PPAs to the engineer, and in return the development of PPAs by the engineers will reflect the interests and concerns. By understanding these interactions between the engineers and groups of interest, the types of decisions and tradeoffs that the engineer makes in the development of PPAs can be realized and allow for the understanding of the social construction of PPAs. Through this SCOT model, the potential future interpretations of PPAs can be approximated and considered, and from this we can come to an understanding on the direction of PPAs as a technology and make a determination if and how the technology will stabilize in such a way that each social group benefits.

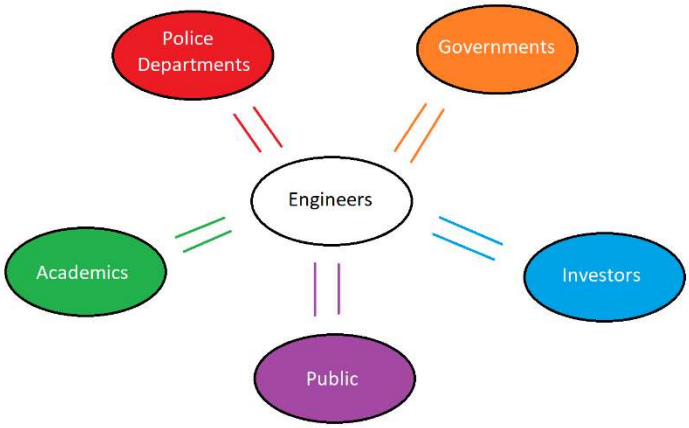


Figure 2: Predictive Policing Algorithm Development SCOT model. Each group and the engineer provide to each other, causing the engineer to make tradeoffs in development to satisfy each group. (Smith, 2021).

Engineers and Academics

As has been the case throughout this paper, academics have been leading much of the discussion around PPAs. Some researchers, such as Shapiro (2017), have called for complete reforms regarding PPAs, whereas others like Yen & Hung (2021) have called for changes while still having an optimistic view of PPAs. The key thing across all academics, though, is that PPAs should not be abandoned; as Yen & Hung put it (2021), many academics criticisms are not inherent to PPAs, but rather their current implementations. Instead, academics and engineers

should continue to iterate and improve on PPAs, as has been done already, which has resulted in perceived benefit to each group. For example, the framework introduced by Lum & Johndrow (2016) has greatly increased PPA awareness of bias and potential discrimination, and given PPAs a path of development to have a minimal amount of bias and discrimination. Similarly, research trials done such as the ones by Ferguson (2017 & 2020) ensure that PPAs are actually effective and provide a measurable benefit. This has allowed engineers to improve on PPAs for increased effectiveness and reception, and in turn the academics are rewarded with accreditation and new prospects for research (PredPol, 2021). Therefore, this relationship is perceivably positive for both groups and will likely lead to PPAs being designed in-line with academic consensus.

Engineers and Investors

One of the hardest relationships to realize, and also quantify, is how the relationship of engineers and investors affect the social construction of PPAs. As pointed out earlier by Selbst (2017), software engineers who design PPAs may inadvertently, or even intentionally design or implement their algorithms in a way which discriminates against certain categorizations of people. This could especially be the case when software engineers may skew their algorithms to deliver the best numerical results to report to their shareholders, and would likely cause for this relationship to negatively impact other groups. However, positive PPA reception is in the best interest of the investors, too, as negative reception has led to governments like the city of Santa Cruz to outright ban PPA use, which directly cuts the profits of the investors (Uberti, 2020). Additionally, with calls for public hearings, government oversight, and increasing transparency by other groups, the negative impact of this relationship towards other groups should be minimized (Yen & Hung, 2021). Ultimately, this relationship is viewed beneficially by both

parties; engineers will continually be funded by the investors, and investors will continually profit from PPAs, and this should cause investors to influence the development of PPAs to be as effective, profitable and well-received as possible.

Engineers and Police Departments

The main benefactors of PPA development, police departments greatly benefit from the implementation and continued development of PPAs as noted by the examples provided by PredPol (2021). Police departments are able to provide feedback towards PPA effectiveness on a near daily basis, allowing the police to essentially be real-time testers for deployments of PPAs that can allow for a constant cycle of development, testing, and refactoring to optimize the effectiveness of PPAs, which shows the relationship greatly benefits the engineers as well. This enables the police to heavily influence PPA development towards more effective and efficient use in crime prevention. Similarly, since police departments draw the majority of their power from the government, and by extension, the public, they will likely always try to ensure that their influence and use of PPAs only serves to benefit these groups as well.

Engineers and Governments

Although not as directly involved in the development of PPAs compared to other groups, governments are closely connected to development via extension to various other groups involved in development. Namely, they must ensure the rights of the public through the control of their institutions, which easily extend to the academics, police, investors and engineers. For example, when the Chicago 'heat list' was determined to be statistically ineffective and discriminatory by Saunders, et al. (2016), they could ensure the rights of the public by cutting funding to Chicago's police department or require reforms. Similarly, if the situation brought about by Selbst (2017) were to come about, where investors and engineers maliciously design

algorithms for profit, they could start reform committees, launch investigations, or ban PPA use from police departments entirely. So whilst they may not influence PPA development directly, indirectly they can actually control all social groups to require PPAs to develop as they see fit. However, major influence by governments on the development on PPAs will likely be rare unless they cross an extreme, and this relationship will seek to ensure all groups benefit, or at least are not negatively impacted, by PPA development.

Engineers and the Public

Despite being the group that is most affected by PPAs, the public has little to no direct say over the development of PPAs. And, currently, the public are the ones facing considerable negative impacts. They are potentially disproportionately affected by PPAs such as in the Chiago 'heat list' case, where some are unfairly put under police surveillance and put at an increased chance of arrest under no basis (Saunders, et al., 2016). Even in the case with Lum & Johndrow's (2016) research, which shows current PPA implementations have demonstrable racial and economic bias towards regular citizens, individuals were not the main ones behind the change to their benefit, but rather the academics were. Fortunately, the public does have some indirect influence over the development of PPAs in that they can enable a government to push their agenda. This indirect influence then tricks down to the other groups, like academics, which then brings forth papers like *Reform Predictive Policing* (Shapiro, 2017) to influence PPA development in their favor. This then leads to changes that benefit them like Santa Cruz banning PPAs (Uberti, 2020) or companies changing their approach to be less discriminatory (PredPol, 2021). Additionally, with academics calling for bringing the development of PPAs in-line with the public agenda via hearings and increased transparency (Yen & Hung, 2021), the public

should gain increasingly more say towards PPAs, which should only serve for them to view PPAs more positively and have more influence over PPA development in the future.

Conclusion

Through the synthesis of academic research and sociotechnical analysis brought forth by this paper, it is easily argued that predictive policing algorithms development is heading increasingly in a direction to better benefit all groups involved that are impacted by its implementation. While there is likely not a future for person-based PPAs, as they only serve to some groups detriment while providing others with no benefit, location-based PPAs have been shown to potentially benefit all groups involved in its development and use. Despite multiple issues with current location-based PPA implementations, which negatively impact some groups involved, these issues have been shown not to be inherent to the technology, with many potential solutions that can and have already been shown to completely remove these issues if not flipping them into something beneficial. In fact, should these issues be resolved, all groups involved with PPAs would have their interests aligned and seek to benefit from PPA use. Therefore, PPAs should continue to be developed, implemented and utilized as a technology moving forward as it would stabilize in a manner that benefits all groups involved.

References

Allen, J. & West, D. (2018, April 24). How artificial intelligence is transforming the world. Retrieved November 1, 2021, from Brookings website: https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/

Bachner, J. (2013). Predictive policing: preventing crime with data and analytics. Washington, DC: IBM, Centre for the Business of Government, 86–90.

Chui, M., Manyika, J., Miremadi, M., Henke, N., Chung, R., Nel, P. & Malhotra, S. (2018, April 17). Notes from the AI frontier: applications and value of deep learning. Retrieved November 1, 2021, from McKinsey website: https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning

Degeling, M. & Berendt, B. (2017, May 22). What is wrong about robocops as consultants? A technology-centric critique of predictive policing. AI & Society, 33, 347-356. https://doi.org/10.1007/s00146-017-0730-7

Ferguson, A. (2017). Policing predictive policing. Washington University Law Review, 94, 1109–88.

Ferguson, A. (2020). Predictive policing theory. Washington College of Law Research, 24, 2020–10.

Mohler, G., et al. (2015). Randomized controlled field trials of predictive policing. Journal Of The American Statistical Association, 110(512), 1399-1411. https://doi.org/10.1080/01621459.2015.1077710

Netherlands: we sense trouble: automated discrimination and mass surveillance in predictive policing in the Netherlands. (2020, September 29). Retrieved November 1, 2021, from Amnesty International website: https://www.amnesty.org/en/documents/eur35/2971/2020/en/

Perry, W., McInnis, B., Price, C., Smith, S. & Hollywood, J. (2013). Predictive policing: forecasting crime for law enforcement. Retrieved November 1, 2021, from RAND Corporation: https://www.rand.org/pubs/research_briefs/RB9735.html

PredPol Overview. (2021). Retrieved November 1, 2021, from Crunchbase website: https://www.crunchbase.com/organization/predpol#section-overview

Rieke, A., Robinson, D. & Yu, H. (2014). Civil rights, big data, and our algorithmic future. Retrieved November 1, 2021, from Big Data Fairness website: https://bigdata.fairness.io/

Saunders, J., Hunt, P. & Hollywood, J. (2016, August 12). Predictions put into practice: a quasi-experimental evaluation of Chicago's predictive policing pilot. Journal of Experimental Criminology, 12, 347-371. https://doi.org/10.1007/s11292-016-9272-0

Selbst, A. D. (2017). Disparate impact in big data policing. Georgia Law Review, 51(1), 109–195.

Smith, J. (2021). Predictive Policing Algorithm SCOT Model. [Figure 4]. STS Research Paper (Societal Impacts of Predictive Policing Algorithms). School of Engineering and Applied Science, University of Virginia. Charlottesville, VA.

Shapiro, A. (2017). Reform predictive policing. Nature, 541(7638), 458-460. https://doi.org/10.1038/541458a

Uberti, D. (2020, June 25). California city bans predictive policing. The Wall Street Journal. https://on.wsj.com/3b9sYIU

Yen, C., & Hung, T. (2021). Achieving equity with predictive policing algorithms: a social safety net perspective. Science And Engineering Ethics, 27(3). https://doi.org/10.1007/s11948-021-00312-x