Norwegian Institute of International Affairs

Policy Brief
[10 / 2018]

Predictive peacekeeping

Opportunities and challenges

Allard Duursma and John Karlsrud

Summary

The time is ripe for the development of a UN early warning tool that estimates the likelihood of instability, intercommunity clashes and armed violence in areas in which UN peacekeepers operate. However, this development would require at least some initial collaboration between the UN and the scientific world. Scientists have developed advanced analytical tools to predict armed violence in recent years.[1] Yet, these conflict prediction tools still cannot be utilized to their full potential because of a relatively poor quality of conflict data. It is precisely in the area of high quality conflict data that the UN has a strong comparative advantage,[2] especially now that the Situational Awareness Geospatial Enterprise (SAGE) system is being implemented. SAGE is a web-based database system that allows UN military, police and civilians in UN peace operations (both UN peacekeeping operations and special political missions) to log incidents, events and activities. The development of SAGE has made it possible to leverage state of the art methodological tools to enable predictive peacekeeping. This policy brief provides background to the recent turn to using data in UN peacekeeping missions, suggestions for what an early warning tool based on SAGE data would look like, and discusses the practical and ethical challenges of such an early warning tool.

Necessity is the mother of invention: current UN data-driven practices

UN peacekeeping has evolved from being considered as an antiquated and outdated organization by western member states to increasingly driving the adoption of new technology at the UN. This trend arguably started when consensus emerged in the early 2000s that the UN should be allowed to produce more efficient field intelligence for its peacekeeping missions. The UN Department of Peacekeeping Operations (DPKO) decided in 2006 that all peacekeeping missions should have a Joint Mission Analysis Centre (JMAC).[3] Recent reports of the Expert Panel on Technology and Innovation in UN Peacekeeping (2014) and the High-level Independent Panel on Peace Operations (HIPPO, 2015) have stressed the role of technology and the latter report also emphasized the need to strengthen the analytical capabilities of peace operations.[4] Following up the HIPPO recommendations, the Secretary-General has tasked the UN Secretariat to develop "parameters for an information and intelligence framework that can support field missions in operating effectively and safely".[5] Recent initiatives in the field have also moved the UN in the right direction. For instance, the UN mission in the Central African Republic (MINUSCA) has developed a Flashpoint Matrix to identify risks for physical violence against civilians, and facilitate a multidimensional response.[6]

However, to become truly data-driven, UN peace operations still have a long way to go. Many of these processes are still ad-hoc and based on local innovation. The rollout of SAGE, which is based on the Ushahidi platform[7], has significantly improved the UN information gathering capacity and is a long step in the right direction. Instead of just reporting free text, the information in SAGE is stored as structured data. This means that the event is categorized (type of event, # of victims, ethnicity, # and affiliation of perpetrators, geographical coordinates and so on). SAGE is an integral and core part of the Mission Common Operational Picture (MCOP), being developed at the time of the publication of this policy brief. Over time, the gathering of structured data will enable mission leadership to identify trends and indicators for early warning.

Hence, while to date much of peacekeeping information gathering efforts have been set up in an ad hoc manner, efforts are currently under way to implement more standardized structures for information gathering within peace missions. This paves the way for the implementation of what we refer to as *predictive peacekeeping*. With predictive peacekeeping, we mean a range of analytic tools and peacekeeping practices that serve to estimate where and when armed violence is likely to take place, combined with changes in peacekeeping leadership decision-making, particularly deployment of peacekeeping staff, based on those estimations. This definition draws on conceptual work on predictive policing aimed at crime prevention.[8] Predictive peacekeeping is thus both about the early identification of a threat and early action aimed at mitigating this threat. The next section focuses on the early identification of threats, discussing the possible use of machine learning to produce a UN early warning tool based on SAGE data.

Early warning 2.0: the possible use of machine learning within peace operations

Machine learning can be defined as "the automated detection of meaningful patterns in data",[9] in other words – learning and making predictions from data. Predicting housing prices serves as a good example of how machine learning could be used to detect meaningful patterns in data inductively (by examples). Housing prices vary according to size and location, but also design, age, access to sunlight, neighborhood, and so on. By feeding a lot of cases where these factors are categorized, we can gradually improve our algorithms to predict housing prices more accurately in an automated fashion. Machine learning is often divided into supervised and unsupervised learning. Supervised learning is when an algorithm is taught the relationship between input data and outputs (called training data), and gradually improve their predictions. Unsupervised learning techniques does not require a training data set and detects patterns in data themselves.

The use of machine learning in UN peacekeeping would mostly be a case of supervised learning, where algorithms are developed, tested and tweaked to constantly improve their predictive capacity. The categories in the SAGE database would be equivalent categories to the housing price example. A major advantage is that machine learning algorithms can take into account how various events and developments *combine* to affect outcomes. For example, a tip-off of an impending attack might not be a significant predictor of armed violence, but in combination with reports on actual troop movements, it might be a highly significant predictor. Machine learning thus makes it possible to grasp the interdependence of all types of incidents reported in SAGE and on other platforms if/when these are interconnected.

Arguably, the most fundamental challenge to predicting armed conflict in space and time will be to obtain high quality data. Conflict processes are incredibly complex because, as noted by Cederman and Weidmann, they "typically encompass an unwieldy set of actors interacting in surprising and, by definition, rule-breaking ways".[10] It is precisely this complex nature of armed conflict that makes it important to not be too overly optimistic of the potential of predicting armed conflict. The first generation of conflict prevention models within academia used to predict the onset of civil wars drew on "sluggish" variables that changed very little from one year to another year like a country's population size. Since these variables do not change a lot, it is hard to predict change (from peace to war) with these variables.

The UN has a comparative advantage in its ability to draw on excellent data. Duursma recently showed that the JMAC conflict data on Darfur is much more comprehensive and precise than the data collected by the Armed Conflict Location & Event Data Project (ACLED), which is widely considered the gold standard when it comes to the collection of conflict data in real time based on media reports. Crucially, JMAC data typically also include observations on variables with a high level of variation – like troop movements of armed actors or tensions identified by informants[11] – which significantly increases the potential for early warning. The SAGE data described in this article holds the potential to be far more comprehensive and precise.

Indeed, the use of peacekeeping data will make it possible to leverage new kinds of predictors that previously have not yet been used when estimating the likelihood of conflict on the subnational level. Cederman and Weidmann warn that machine learning based on big data that is not "theoretically informed" will probably not significantly improve conflict prediction models: "Ultimately, the hope that big data will somehow yield valid forecasts through theory-free 'brute force' is misplaced in the area of political violence. Automated data extraction algorithms, such as Web scraping and signal detection based on social media, may be able to pick up heightened political tension, but this does not mean that these algorithms are able to forecast low-probability conflict events with high temporal and spatial accuracy". Peacekeeping data, by contrast, typically includes observations that, from a theoretical perspective, should be strong predictors of armed violence on the subnational level.

Consider the following example of an event from the JMAC dataset on Darfur that goes well beyond observations that are typically used to predict armed violence on the subnational level like armed clashes, violence against civilians, and protests: "*On 20 May 08, SLA/MM stated that according to their information GoS affiliated armed groups are mobilizing with the purpose to attack several SLA/MM controlled villages, including Muhajeriya, Labado and Marla*".[13] Leveraging these kinds of predictors for estimating the likelihood of armed conflict has great potential. In addition, UN information analyst usually have a good understanding of – as for example insights where roadblocks are most likely to be placed (e.g. bridges and junctions). This type of information can also be included in early warning models.

Practical implications: from early warning to early action?

Using SAGE data for machine learning can offer a giant step forward in the predictive capability of the UN, and hopefully be translated to preventive action on the ground. Predictive analyses could take the form of at least two specific outputs. A first output could be a risk map in which a color coding of administrative districts indicates the probability of events of interest like armed clashes between the main warring parties, communal violence, or violence against civilians.

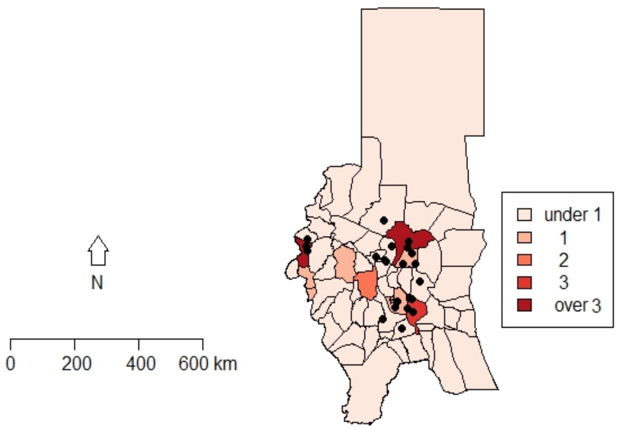

Figure 1: The Spatial Correlation between JMAC Reports on Tensions in October, November, and December 2008 and Violence against Civilians in Darfur in January 2008

The map above in Figure 1 graphically illustrate how the number of tensions reported in localities in Darfur between October and December 2008 are spatially correlated with the number of attacks on civilians in January 2009.[14] The localities with the light shade were covered in the JMAC dataset on Darfur as having experienced no tensions in the previous three months. The darker the colour of the locality, the more tensions were reported in the locality in the previous three months. The black bullets plotted on the map are instances of violence against civilians in Darfur in January 2009, from which it follows that violence against civilians predominantly occurred in those areas which have covered in the JMAC dataset as having experienced tensions.

It should be noted that this map only considers the number of tensions reported in a given locality and for an extremely limited time frame. The use of machine learning based on SAGE data will make it possible to calculate the predicted probabilities of armed violence in a much more precisely defined area, on a range of different variables, and over a relatively long time frame. Hence, the predictions made based on SAGE data are likely to be much more accurate compared to the identification of violence against civilians as done in in Figure 1.

A second output could be a user-friendly alert tab function in the SAGE interface where UN staff can create alerts using the categories and indicating a frequency of such violent incidents in a given time period. Based on these inputs, an alert would be generated when the combination of violent incidents may signify the imminent unfolding of a larger conflict event of interest to them (e.g. communal violence) This alert could be given with a statistical plausibility level attached. For instance, using the alert tab in the SAGE interface, a United Nations Mission in South Sudan (UNMISS) staff member could sign up for receiving an alert message as soon as the *predicted* number of incidents involving communal violence in Ruweng State in South Sudan is more than three for the next week. As soon as the machine learning algorithm predicts that the level of communal violence in Ruweng State indeed hits this threshold, an email would be sent to the person for further evaluation and action – e.g. sending more frequent patrols or redeploying a company of the quick reaction force, and sending a Civil Affairs team to the locale to get a better understanding of the situation and facilitate talks between the parties involved.

Ethical considerations

The data lifecycle contains numerous and significant risks. The collection and storage of sensitive data necessitates increasingly strong rules and routines for management of data and information. How and for how long will the information be stored, who will have access to it, and what types of security measures will be taken at all levels to ensure the integrity and safety of the data? The UN has already become the target of offensive cyber-attacks. Attacks could aim to retrieve data or even change data to alter the understanding of the reality on the ground.

Privacy concerns will also be central – civilians who are already at risk can face new threats if their personal information is disclosed or reidentified. A telling example in this regard is the confidential relationship UN information analysts have with local informants. Local informants often play a crucial role in the day to day work of information analysts, as they can provide information that is very hard to get from other sources.[15] However, this valuable information comes at a price, as conflict parties may target individuals that are suspected to have passed on information to the UN in order to dissuade potential other informants from also sharing information. If, for some reason, UN data on local informants is leaked, then the consequences could be fatal.

Technological advances improve the ability to understand the operational situation from afar, but this also increases the risk of remote management. Remote management can reduce the UN's ability to interact, understand and empathize with local populations, who, after all, are those UN peace operations should be most accountable to. Through innovation and simulation, technology can replace ground truth, adding ammunition to those who criticize the UN for an increasing tendency of "bunkerization" – retreating behind the safe confines of high walls and Hesco barriers.[16]

Finally, obtaining early warnings also creates expectations for the UN to act. Edward C. Luck, the Special Adviser to Secretary-General Ban Ki Moon, has pointed out in this regard hat early warning is not an end in itself: 'Early warning without early and effective action would only serve to reinforce stereotypes of UN fecklessness, of its penchant for words over deeds.'[17] An improved early warning capacity should therefore be accompanied with a clear policy on how to turn early warning into early action.

Conclusion

The data which is being gathered, categorized and stored in SAGE can be analyzed using machine learning techniques. This is a positive development, enabling preventive deployment to protect civilians and staff alike. Leveraging SAGE data using machine learning techniques will probably make it possible to identify patterns that currently are not apparent. This could lead to new and innovative methods and tools for protection. Different types of interventions could be tested, including military, police, and civilian patrols, as well as the use of surveillance UAVs. However, it will require resources as well as careful thinking about the potential pitfalls that practitioners and policy makers will be confronted with.

The UN and the scientific community thus need to consider the implications of current and future developments in this area. What are current and future ethical challenges that will be faced? How can it be ensured that improved information indeed leads to improved action? Where and in what tools should member states invest to support these technological developments in the best possible manner? This policy brief has sketched out tentative answers to some of these questions, but much work remains.

Endnotes

1. Perry C. (2013) Machine Learning and Conflict Prediction: A Use Case. *Stability: International Journal of Security & Development* 2; Chadefaux T. (2014) Early warning signals for war in the news. *Journal of Peace Research* 51: 5-18; Muchlinski D, Siroky D, He J, et al. (2017) Comparing Random Forest with Logistic Regression for Predicting Class-Imbalanced Civil War Onset Data. *Political Analysis* 24: 87-103; Blair RA, Blattman C and Hartman A. (2017) Predicting local violence: Evidence from a panel survey in Liberia. *Journal of Peace Research* 54: 298-312; Colaresi M and Mahmood Z. (2017) Do the robot: Lessons from machine learning to improve conflict forecasting. *Journal of Peace Research* 54: 193-214.

2. Karlsrud J. (2014) Peacekeeping 4.0: Harnessing the Potential of Big Data, Social Media and Cyber Technology. In: Kremer JF and Müller B (eds) *Cyber Space and International Relations. Theory, Prospects and Challenges.* Berlin: Springer, 141-160; Duursma A. (2017) Counting Deaths While Keeping Peace: An Assessment of the JMAC's Field Information and Analysis Capacity in Darfur. *International Peacekeeping* 24: 823-847.

3.Shetler-Jones P. (2008) Intelligence in Integrated UN Peacekeeping Missions: The Joint Mission Analysis Centre. *International Peacekeeping* 15: 517-527.

4. UN (2014) *Performance* Peacekeeping. New York: United Nations; UN (2015) *Uniting our strengths for peace: politics, partnership and people*. New York: United Nations.

5. Ibid.: 94.

6. O'Bryan T, Rendtorff-Smith S and Donati M. (2017) The Role of United Nations Peacekeeping Operations in Addressing Local Conflicts: A Study of Practice. *United Nations Departments of Peacekeeping Operations and Field Support Policy, Evaluation and Training Division Policy and Best Practices Service*.

7. Ushahidi, which translates to "testimony" in Swahili, was developed to map violent incidents in Kenya after the post-election violence in 2008.

8. See: Uchida C. (2014) Predictive policing. In: G Bruinsma and D Weisburd (eds.) *Encyclopedia of criminology and criminal justice*. New York: Springer: 3871.

9. Shalev-Shwartz S and Ben-David S. (2014) *Understanding Machine Learning: From Theory to Algorithms*, Cambridge: Cambridge University Press: vii.

10. Cederman L-E and Weidmann NB. (2017) 'Predicting armed conflict: Time to adjust our expectations?' *Science* 355: 475.

11. Duursma, 'Counting Deaths While Keeping Peace'.

12. Cederman and Weidmann, 'Predicting armed conflict': 476

13.Duursma, 'Counting Deaths While Keeping Peace'.

14. Taken from Duursma, 'Counting Deaths While Keeping Peace'.

15. Duursma A. (2018) Information Processing Challenges in Peacekeeping Operations: A Case Study on Peacekeeping Information Collection Efforts in Mali. *International Peacekeeping* 25: 446-468.

16. Duffield M. (2013) *Disaster Resilience in the Network Age: Access-Denial and the Rise of Cyber-Humanitarianism*, Copenhagen: Danish Institute for International Studies.

17. Luck, E. C. (2010) Statement by Edward C. Luck Special Adviser to the United Nations Secretary General Informal Interactive Dialogue on Early Warning, Assessment, and the Responsibility to Protect United Nations. United Nations General Assembly 9 August 2010.

Norwegian Institute of International Affairs

NUPI
Norwegian Institute of International Affairs
C.J. Hambros plass 2D
PO Box 8159 Dep. NO-0033 Oslo, Norway
www.nupi.no | info@nupi.no

Established in 1959, the Norwegian Institute of International Affairs [NUPI] is a leading independent research institute on international politics and areas of relevance to Norwegian foreign policy. Formally under the Ministry of Education and Research, NUPI nevertheless operates as an independent, non-political instance in all its professional activities. Research undertaken at NUPI ranges from short-term applied research to more long-term basic research.

Allard Duursma is a Senior Research Fellow at the Center for Security Studies (CSS) at ETH Zürich where he works on early warning, international mediation and peacekeeping.

John Karlsrud is a Senior Research Fellow at NUPI where he works on peacekeeping, peacebuilding and humanitarian issues, with a particular interest in how technology and data can be leveraged to make peace operations more effective and accountable both horizontally and vertically.

Acknowledgments

The authors are grateful for comments and inputs received from colleagues at NUPI, and participants at the International Studies Annual Convention 2018 and the Ending Atrocities Workshop at Uppsala University in June 2018.