Grey Zone activity: measuring the resilience of social systems to influence operations

Marian-Andrei Rizoiu, *University of Technology Sydney*
Thomas Willingham, *Australian National University*
David Kernot, *Defence Science and Technology Group, Department of Defence*

Hostile foreign actors are upgrading their suite of cyber tools along a hybrid war spectrum that deny critical infrastructure and industries but also seek to disrupt **human systems**. These grey-zone activities [1] occur with little or no warning below the threshold of armed conflict. While the strength of communication systems is a relatively well-studied domain, this is less true of social systems.

Our work studies the resilience of information and social systems in contested environments, using a blend of modern Artificial Intelligence and psycholinguistics. We analysed the discussions on online social mediums around controversial topics [2] – exactly the type of topics susceptible to information and influence operations. We followed topical Australian conversations – more than 17 Million tweets centred around the Australian 2019 election, the Q+A talk show, and political issues like StopAdani and ClimateChange. We studied the societal responses in the face of disinformation spread by developing tools that combine Network analysis, Machine Learning and Artificial intelligence with two main goals: a) to measure the societal response, and b) to understand the flow and spread of disinformation, widely used in influence operations. We found evidence that defined groups wielded a constant barrage of information, whose timing and linguistic patterns are indicative of coordinated non-organic actions. We also quantified emotions, grievances, and other psycholinguistic variables to create a measure of community resilience to such actions that can act as a societal barometer.

Our tools and research can support existing information warfare capabilities, by measuring the resilience of human systems to influence operations [3], and provide the means to construct counternarratives. Our technology can be readily integrated by the Australian Defence Force and other government agencies to support information warfare capabilities to monitor, detect and counter influence operations, defend contested information environments, and support community resilience during bushfires and other natural disasters.

References:

[1] Pettersson, D., Popkewitz, T. S., & Lindblad, S. (2017). In the grey zone: large-scale assessment-based activities betwixt and between policy, research and practice. *Nordic Journal of Studies in Educational Policy*, 3(1), 29-41.

[2] Kong, Q., Booth, E., Bailo, F., Johns, A., & Rizoiu, M.-A. (2022). Slipping to the Extreme: A Mixed Method to Explain How Extreme Opinions Infiltrate Online Discussions. *Proceedings of the International AAAI Conference on Web and Social Media*, *16*(1), 524–535. https://doi.org/10.1609/icwsm.v16i1.19312

[3] Kong, Q., Calderon, P., Ram, R., Boichak, O., & Rizoiu, M.-A. (2023). Interval-censored Transformer Hawkes: Detecting Information Operations using the Reaction of Social Systems. In *Proceedings of the ACM Web Conference 2023* (pp. 1813–1821). New York, NY, USA: ACM. https://doi.org/10.1145/3543507.3583481