Machine Reasoning for Determination of Threat Level in Irregular Warfare

Charles Kim
Electrical Engineering and Computer Science
Howard University
ckim@howard.edu

Summary

The objective of the research is development of a machine-reasoning based decision-assist system for determining and predicting threats for field commanders in irregular warfare. Centered on information entropy and Bayesian inference, the machine reasoning system extracts dominant contributory attributes to threat and generates decision rule for the threat and the certainty level of the rule itself. The reasoning system utilizes diverse data of all available, local and global, human terrain resources.

Abstract

Irregular warfare (IW) campaigns depend on not just military prowess but also understanding of social dynamics. Therefore IW or counterinsurgency cannot be conducted without understanding the human terrain. The objective of the research is development of a machine-reasoning based decision-assist system for determining and predicting threats for field commanders in irregular warfare utilizing diverse datasets of local populace environment and other sources from the other parts of the world. The machine reasoning system aims to answer the following question: With the present information, what is the threat probability of the region of interest and how certain is the probability itself?

Provision of quantitative threat level given diverse datasets and information requires an intelligent system which extracts dominant contributors and learns and updates as new data is added to the datasets. The machine reasoning system keeps, upon existing and updated datasets, extracting dominant contributory HT attributes, generating rules for threat level determination with the attributes, and producing the certainty level of the rules themselves.

The main theory behind dominant attribute discovery and decision rule extraction from datasets is the information entropy minimum principle. “Information measure” (I) is defined as proportional to the negative of the logarithm of probability (p), with k a constant: $I = -k*\ln(p)$. Information entropy (S) is defined as the expected value of information: $S= -k*p*\ln(p)$. In the entropy minimum state, all of the information has been extracted, and there is no information gain, leading to maximum certainty.

The first step in determining the dominant attribute is to convert all analog-valued sample data to binary valued data. The “binarization” is performed by threshold calculation. The calculated threshold value with minimum conditional entropy optimizes the separation of two outcomes, Threat (*T*) and No-Threat (*F*). The conditional entropy $S(x)$, which, for a chosen value x, is defined with conditional probabilities of two outcomes, T and F, under 2 conditions (one for a sample value lower (x-) than a certain threshold value x and the other greater (x+) than that) as,

$$S(x) = -p(x-)\,[p(T|x-)\ln(p(T|x-)) + p(F|x-)\ln(p(F|x-)) - p(x+)\,[p(T|x+)\ln(p(T|x+)) + p(F|x+)\ln(p(F|x+)).$$

A binarized sample data is obtained after converting the analog data into binary values, 1 for sample values above the threshold, 0 for below the threshold. The same conditional entropy can be applied to

determine dominant attributes in correlating an attribute to the outcomes. A conditional entropy equation for the *i*th attribute, Si, for T or F under 0 or 1 attribute value is as follows:

Si = -pi(0) [pi(T|0)ln(pi(T|0)) + pi(F|0)ln(pi(F|0)) -pi(1) [pi(T|1)ln(pi(T|1)) + pi(F|1)ln(pi(F|1)).

After applying the conditional entropy to all *m* attributes, a certain attribute Ak which produces the minimum conditional entropy will be the best attribute in correlating the sample data to the outcomes. Then the decision rule, R_k for the attribute *k*, can be drawn from the best (highest) conditional probability from the set of four: $p_k(T|1)$, $p_k(F|1)$, $p_k(T|0)$, and $p_k(F|0)$.

If, for example, $p_k(T|1)$ is the highest from the set, then the decision rule is formed as follows:
R_k: IF (A_k = 1), THEN (T).

In this step, the probability (or certainty) of this decision rule itself is generated from the maximum entropy based Bayes estimate by <p(O)> = {x + 1 }/ {n + 2}, where, x is the total number of samples satisfying the condition (T|1), and n is the total number of samples satisfying the attribute condition. Also, the margin of error of the drawn probability is obtained by e(O)= z*√[{ <p(O)>* (1 – <p(O)>}/{n+2)] ,where z is a z-score value for desired confidence interval.

Usually, not all samples can be directly linked to a single decision rule. Therefore, we apply step-wise approximation by which, after the first attribute and its corresponding decision rule are found, we remove all the samples which match the decision from the binarized dataset and we repeat the conditional entropy minimum process for the remaining data samples.

We tested the implemented machine reasoning system with an example dataset which has a total of 31 samples, 21 Threat and 10 No-Threat samples, with 7 attributes. The reason produced 3 steps of rules with a satisfactory result.

For practical applications, the decision assist platform would be configured around a client-server based secure network. The machine reasoning algorithm resides in a server and the data would be loaded to the server from client computers in the network. A client representing a field command console for threat level monitoring would also be connected to the network. The machine reasoning system would provide the forces in the field with operationally-relevant decision assistant for determining threat in global context.

Biography

Dr. Charles Kim is a professor in Electrical Engineering and Computer Science at Howard University. He has experienced in applying information entropy and Bayesian inference as an intelligent and machine learning system to time-series data for diagnostic detection, classification, and pattern discovery. He has applied the machine reasoning to deal with electronic component assessment for radiation impact and sensitivity during his summer research at NASA Jet Propulsion Laboratory in 2017. Recently, the application area has further extended to social science and human-terrain related problems for decision-making and behavior classification. His present research is focused on threat level determination in the irregular warfare extracted from local and global data of human terrain. Dr. Kim earned his Ph.D. in electrical Engineering at Texas A&M University.