A View on Threat Analysis Concepts, Models and Estimation Techniques

Jean Roy
Defence R&D Canada – Valcartier
2459 Pie-XI Blvd. North, Quebec (Quebec)
Canada, G3J 1X5
jean.roy@drdc-rddc.gc.ca

Abstract - *This position paper summarizes a recent attempt by the author to synthesize what recognized experts in the information fusion and related domains have previously said about threat analysis, while at the same time injecting the point of view of the author. Aspects of modeling the behaviour of humans considered as limited, rational purposeful agents are first presented. Then, the notions of impact and threat as « things » resulting from actions or, more precisely, from the state changes caused by actions are discussed. The important concept of a concern reference point is introduced and described. The three essential and interdependent parts of a threat, i.e., intent, capability and opportunity are briefly presented. Finally, a definition of threat analysis is proposed, and some techniques that have been used for threat and intent estimation and prediction, for behaviour analysis, and for plan analysis and recognition are listed.*

Keywords: Threat Analysis, Impact Assessment

1 Modeling the behaviour of an agent

It is convenient, for the purpose of threat analysis, to consider humans as agents that perform intentional actions, making use of the reason, and taking into account constraints imposed by their limited resources [1]. Humans have « sensors » and they observe the environment (including the other humans). They acquire knowledge and develop awareness of the situations. Their sensors having limitations, there is uncertainty about their perception, understanding and prediction of these situations, and they express their views through beliefs. They develop desires (or high-level goals), and they have resources and capabilities to fulfill them. They make decisions and formulate intents about some end states of the world. They have natural needs for coordination and planning to guide their behaviour, i.e., their actions.

In general, actions can be thought of as involving one or more actors (or "agents") and one or more "objects" acted upon (e.g., other agents or the environment) [2]. Typical situations of interest involve some interactions between the action plans of multiple agents [3]. Ultimately, a given state change resulting from an action may have a large impact on the plan of another agent, forcing this agent to cancel upcoming planned actions. To recognize and anticipate other agents' actions, a node in a network of interacting agents must model such agents' actions in response to a world state or world state history [4]. A model for the adaptive response of a limited, purposeful and rational human agent has been proposed in [1].

2 Impact/threat ensuing from actions

At the heart of the "behaviour process" model described in [1] is the notion of an action. The actions potentially leading to threat events may be intentional or unintentional. Examples of typical actions of interest are an attack, an intrusion, a warning, the capture of someone or something, the surveillance of an area, a strike, a deployment, the blast of a bomb, etc. People typically have the power to control or direct themselves, i.e., to manage their actions in a particular way. This leads to behaviour, which is often characterized as normal or suspicious/abnormal. Actions create some activity and courses of events. All actions performed by an agent perturb the environment, i.e., they produce some alterations of the state of the environment (potentially including the state of other agents). Actions, and mostly their resulting state changes, play a key, central role in any discussion on impact assessment and threat analysis. As situations are dynamic, there are the related notions of current state, future state, end state, etc.

The analysis of the nature and magnitude of the state changes resulting from some action (and also those of the related effects/consequences/impacts) clearly depends on the perspective or context being considered, i.e., on some *concern reference point* (CRP). Moreover, the state changes and related consequences may be considered as being "negative" or "positive", depending on the particular CRP. Some consequences considered as positive for someone could be considered as highly negative for someone else. Of course, negative consequences are of particular interest in the context of threat analysis, as the notion of a threat has an intrinsic connotation of negativeness.

The term *point* may be used to mean a non-material *point of view*. However, it is often used to refer to the geometrical aspect of the concern perspective, as the nature and magnitude of some state changes and/or consequences often depend on the geometry (e.g., the proximity) between the "effector" (e.g., an agent performing some action) and

the "effected" (e.g., a particular asset). The notion of a CRP can also be linked to the notion of the target of some intentional action(s). Different assets can be the targets of some intentional actions. High-value assets are of particular importance, but the value of an asset clearly depends on the different CRPs that may exist. There is a need to define some taxonomy of targets (and CRPs), because intent, capability and opportunity depend on the type of target being considered. As the "robustness" of the target increases, so do the required capabilities to affect it. As the "importance" of the target increases, so does the adversarial intent to affect it. Opportunities may also be dependent on the nature of the target. Another aspect that can be related to the value of an asset (target) is the payoff, or utility, of performing an action (against it). A purposeful agent typically has expectations regarding the benefit(s) of some action(s) about to be carried out. Usually, the higher the value of the target, the higher the benefit expectations of the agent. The same can be said of the motivation of the agent to reach his/her goal in proportion to the expected payoff.

Once a particular CRP is defined, one can analyse the impact of some actions with respect to this CRP. To put it briefly, impact has to do with what something means with respect to some CRP. As impact (assessment) is also the foundation of threat (analysis), it is important to develop a sound understanding of it. Consequences/effects/impacts have different attributes such as their nature, strength, polarity, intentionality, continuity, manifestation, duration and evolution. Many of these attributes depend on the perspective (i.e., the CRP) of the agent that is evaluating the situation. Examples of impacts/consequences/effects include physical, behavioural, psychological, cyber, social and economical. While nouns are often used to describe their nature, verbs are also often used (modify, disable, influence, dissuade, deter, compel, preclude, damage, disrupt, kill, demoralize, paralyze, slow, divert, confuse, negate, destroy, injure, harm, hurt, etc.), emphasizing the idea that impact is a relation between two things.

One action can produce multiple state changes and, consequently, multiple different effects or impacts, each one depending on a particular CRP. However, the opposite is also possible, i.e., multiple actions can produce a single consequence. Sometime, one is also faced with some interdependency of consequences. There may be a "domino" or cascading effect, some chaining of consequences resulting from a single action or event.

Threats are the objects of agent-based complex activities, which result in some "negative" impact(s). Such negative effects or consequences can be on objects (e.g., some high value assets), or on other agents, including their "behaviour process" (i.e., the other agents' desires, decisions, intents, plans, capabilities, opportunities, actions, etc.), or on political structures, financial institutions, etc. The notion of a threat also has to do with something that is a source of danger, a warning that something unpleasant is imminent, a declaration of an intention or a determination to inflict harm on another, a person who inspires fear or dread, inflicting evil, injury, or damage, the potential for destruction, removal, modification or interruption of some (usually valuable) processes or assets, and the combination of the presence of a hazard and an exposure pathway. Threats are often categorized with respect to the means used to inflict evil, injury, or damage (e.g., chemical, biological, radiological, nuclear threats), or with respect to the type of targeted entities (e.g., cyber, economical threats).

3 Intent, capability and opportunity

An intention is a determination to act in a certain way. In the context of threat analysis, the term intent (implying here hostile or malicious intent) often refers to the will or determination of an entity (or more correctly the human actor(s) controlling the entity) to inflict evil, injury, or damage to the defending forces and their interests. As intent is an intangible item that cannot be directly observed with sensors, it must rather be inferred from the observation of other aspects of the world that becomes indicators of intent. In turn, this requires a model of intent that is sufficiently precise and detailed, providing a fair understanding of the numerous factors that drive intent. Previous efforts towards the definition of a model for intent have included a military perspective on commander's intent, a belief-desire-intent framework, planning-based models of intent, and explicit and implicit intent.

Although it is an important factor regarding one's determination to act, just having a *desire* is not enough to conduct an activity [3]. One must also have the basic capability required to achieve a goal. The notion of capability has to do with the potential ability to do work, perform a function or mission, achieve an objective, or provide a service. It is somewhat linked to the notion of feasibility.

An opportunity is a favourable juncture of circumstances; a good chance for advancement or progress. An opportunity model concerns the presence of situational factors necessary for actions to occur [4]. Opportunities makes it possible to actualize (i.e., carry out) one's intent given sufficient capabilities [5]. An opportunity is the presence of an operating environment in which potential targets of an action are present and are susceptible to being acted upon [4]. The notion of opportunity is highly linked to the notion of target exposition, which has to do with depriving of shelter or protection, subjecting to risk from a harmful action or condition, making accessible to a particular action or influence, or causing to be visible or open to view. Vulnerability is some sort of "mirror concept" to that of opportunity, in the sense that the vulnerability of an entity may become an opportunity for some agent.

Threat items must be understood as wholes containing parts that exist in concert [5]. Viable threats form tri-partite integral whole, which possesses three essential and interdependent parts: intent, capability and opportunity. If these are understood as being connected to one another via relations of dependence, then disruption or dissolution of

any one of these elements results in a disruption or dissolution of the threat as a whole. Viable threats exist when all three essential elements are present. Potential threats exist when at least one essential part, and its corresponding relations, is missing. They are threats that are not in a state of being; they are in a state of becoming.

4 Threat analysis

In view of the discussion above, threat analysis can be defined as « *The analysis of the past, present and expected actions of external entities, covering the overall behaviour process of these entities from desires to effects/consequences, to identify menacing situations and quantitatively establish the degree of negativeness of their impact on the state and/or behaviour process of some agent, and/or on some valuable human/material assets to be protected, taking into account the defensive actions that could be performed to reduce, avoid or eliminate the identified menace.* » The "degree of negativeness" has been included in the definition to reflect the fact that the notion of a threat evokes and involves only negative connotations such as danger, harm, evil, injury, damage, hazard, destruction, loss, unpleasant, fear, dread, etc.

The concept of *inherent threat assessment*, introduced in [3], has to do with quantitatively establishing the degree of (negative) impact of each upcoming consequence resulting from some action(s) performed by the players/actors/agents. The idea is to quantify the intrinsic level of danger or menace (i.e., the potential for causing victims, harm, damage, mission failure, etc.) of each consequence if nothing is done to prevent this consequence from happening. At this stage of the analysis, a friendly weapon's ability, or inability, to engage an entity should not be used to classify this entity as a threat or non-threat.

A process (mathematical and/or rule-based, etc.) assigns *inherent threat values* to consequences. These are numerical values reflecting the degree of inherent threat evaluated according to the estimated impact of some state change(s) resulting from actions. For example, it could be a number between 0 (non threat) and 1 (highest threat value). Getting an understanding of the consequences of the various actions performed (or to be performed) by the various agents in the environment is a critical step of threat value calculation. One is concerned with identifying the nature of each consequence resulting from adversarial actions and events (such as an explosion, a physical hit, the release of some highly toxic chemical substances, etc.), and quantifying the impact of such events on the mission, the intent, the plans, the actions, etc. of some agent(s), or on some valuable human/material asset(s) to be protected.

The second notion introduced in [3] is *actual risk assessment*, which is about taking into account the defensive actions that could be performed to reduce, avoid or eliminate each menace previously identified through the inherent threat assessment process. It has to do with quantifying how easy it is to avoid/defeat each individual threat on some prioritized threat list [3]. One attempts to answer questions such as: Does one know how to tackle the problem posed by the threat? How many defensive options does one have to avoid/defeat the threat? Based on the answers obtained, a *threat value* is transformed into an *actual risk value* that better reflects the actual, tangible potential for danger [3]. An entity that has been assigned a very high inherent threat value could ultimately represent a very small risk if it is very easy to take care of it (e.g., there are numerous, good quality options to tackle the problem). A moderate threat entity may represent a high risk if there are no options available to counter it.

4.1 Estimation and prediction techniques

Some quantitative, numerically-based techniques that have been used for threat estimation and prediction, as reported in the literature, include Bayesian networks, value-based systems, evidential networks, influence diagrams, game theory, Bayesian games, hierarchical task networks, and genetic algorithms. Symbolic techniques, such as rule-based systems, have also been used. Techniques that have been used for intent estimation and prediction include plan-goal graph approaches, blackboards, potential field models of environmental influences, and Bayesian networks. Techniques used for behaviour analysis include rules and fuzzy sets, genetic algorithms, and use case template-based techniques. And finally, techniques used for plan analysis and recognition include Bayesian networks, hidden Markov models for multi-agent plan recognition, and symbolic reasoning with models of adversarial plans.

References

[1] Roy, J., *Modeling the Behaviour and Estimating the Intent of a Rational Purposeful Agent*, DRDC Valcartier Technical Report, TR 2008-381, 2009.

[2] Steinberg, A.N., *Open Networks: Generalized Multi-Sensor Characterization*, Proceedings of the 9th International Conference on Information Fusion (Fusion 2006), Florence, Italy, 10-13 July 2006.

[3] Roy, J., *A View on Threat Analysis Concepts, Models and Estimation Techniques*, DRDC Valcartier Technical Report, TR 2008-382, 2009.

[4] Steinberg, A., *Predictive Modeling of Interacting Agents*, Proceedings of the 10th International Conference on Information Fusion (Fusion 2007), Quebec, Canada, July 2007.

[5] Little, E.G. and Rogova, G.L., *An Ontological Analysis of Threat and Vulnerability*, Proceedings of the 9th International Conference on Information Fusion (Fusion 2006), Florence, Italy, 10-13 July 2006.