WG 3: International Safeguards

Topic: **Developing Metrics for the Detection and Evaluation of Undeclared Nuclear Activities**

Irmgard Niemeyer, Forschungszentrum Jülich GmbH, Germany, i.niemeyer@fz-juelich.de
Arnold Rezniczek, UBA GmbH, Germany, Rezniczek@uba-gmbh.de
Gotthard Stein, Consultant, Germany, g.stein@fz-juelich.de

The acquisition path analysis model, as suggested by Listner at al.[1], can be used to determine an optimal set of technical objectives. As input on behalf of the State, the model requires an assessment of each process' attractiveness and the resulting "payoff" values for each path. On behalf of the inspectorate, for each technical objective a cost estimate is needed as well as an estimate of the non-detection probability for each process given that a specific technical objective is in place. It turns out that the attractiveness values, the cost estimates and the non-detection probabilities in terms of declared facilities can be obtained relatively easily because there are models available for the estimation of these parameters. However, until now there are no models available for the estimation of non-detection probabilities for processes in covert facilities as well as undeclared import.

In the past, the estimation of such non-detection probabilities has been considered to be an impossible task. The reasons for the reluctance to quantify these parameters can be found in the lack of system boundaries of clandestine nuclear facilities as they can be located anywhere in a State. The same applies to the case of undeclared import, where the location of possible indicators could even be found worldwide. Moreover, it is not even clear which indicators could give the relevant hint to a clandestine facility. All these problems seem to be good reasons to tackle the detection of clandestine facilities and undeclared import only in a qualitative way. However, this would lead to the problem of how to justify the budget expenditures on the detection of clandestine facilities against conventional safeguards measures whose effectiveness can be quantified very elegantly. A model calculating quantitative estimates for the non-detection probabilities can overcome this issue. Also, this problem is similar to effectiveness quantification in the intelligence realm and there has been research on how to address this.[2].

In the past it has been shown that hypothesis testing is a powerful tool that can be applied in the context of treaty verification to estimate the errors.[3] It assumes that a State can behave either compliantly or not. In contrast, the inspectorate has the possibility to raise an alarm or not. Thus, the error model results in four event combinations (see Table 1).

	No Alarm	**Alarm**
Compliance	$1-\alpha$	α
Non-Compliance	β	$1-\beta$

Table 1. Verification Error Matrix

The main diagonal entries of this matrix indicate a properly working verification system which raises an alarm in case of non-compliance or does not in case of compliance. The off-diagonal elements however reflect errors in the verification system. An error of the first kind, α, also known as a false alarm, will occur, if the State behaves compliantly but the inspectorate raises an alarm despite that fact. The error of the second kind, β, is also known as non-detection of incompliance. This error will occur, if the State proliferates but the inspectorate is not able to detect this behavior and thus will not raise an alarm.

Based on this error model, the existing literature and some new ideas, four possibilities will be

[1] Listner, C., Murphy, C.L., Canty, M.J., Stein, G., Reznicezek, A. & I. Niemeyer (2015): Acquisition Path Analysis Quantified - Shaping the Success of the IAEA's State-level Concept, *JNMM* (in print)

[2] Lehner, P., Michelson, A. & L. Adelman (2010): Measuring the Forecast Accuracy of Intelligence Products. Tech. rep. MITRE

[3] Avenhaus, R. & M.J. Canty (1996): *Compliance quantified*. Cambridge University Press, Cambridge

presented how to estimate the non-detection probabilities in case of undeclared facilities or import. These suggestions should be seen as a starting point for further discussion and research.

1) The Analogy Approach

The first and by far the simplest possibility starts by looking into declared facilities. There, the safeguards system can obtain a non-detection probability of $\beta_{declared} = 10\%$ if all measures, like e.g. PIVs and IIVs, are in place. By analogy, the same non-detection probability of $\beta_{undeclared} = 10\%$ is assumed for undeclared facilities in case all measures, like e.g. open source information analysis taskings, are applied here as well. If only parts of the measures are applied, a linear scaling procedure increases the non-detection probability.

E.g. in case only half of the measures are applied, the detection probability reduces from $1 - \beta_{undeclared} = 90\%$ to $1 - \beta_{undeclared} = 45\%$.

2) The Bayesian Approach

The second approach uses Bayes' theorem to model the information analysis process and then estimates the detection probability from a simulation step. In this context, the event A_j means that the proliferation activity *j*, e.g. the use of a clandestine reprocessing facility, is carried out by the State. $B = \{B_1, \ldots, B_n\}$ represents the set of available information pieces. Based on these probabilistic events, the Bayes formula retrieves the probability of a proliferation activity A_j given a set of available information B as

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{P(B|A_j)P(A_j) + P(B|\overline{A}_j)P(\overline{A}_j)}$$

Once the Bayes formula is applied to derive the probability $P(A_j|B)$, the information analysis process would raise an alarm, if this probability exceeds a given threshold T. In order to derive the non-detection probabilities β, one checks the correctness of the information analysis process for any combination of B, A_j and $\overline{A}_j$ weighted by the probability of each event combination. The error of the first second kind then gives the non-detection probability β.

3) The Frequentist Approach

As a third possibility, historical events in the field of non-proliferation can be used to retrieve estimates for the non-detection probability. Therefore, the error matrix (Table 1) is filled with the absolute number of events (see Table 4). Using these figures, the false alarm probability non-detection probability can be estimated using

$$\hat{\alpha} = \frac{H_{false\ alarm}}{H_{false\ alarm} + H_{compliance\ without\ alarm}} \qquad \hat{\beta} = \frac{H_{undetected\ non-compliance}}{H_{undetected\ non-compliance} + H_{succesful\ detection}}$$

4) The Process Approach

Finally, the fourth approach considers α and β to be "measurement errors" of the inspectorate's information analysis process. This information analysis process can be subdivided into five components according to the intelligence cycle[4]: plan, collect, process, analyze, disseminate.

For each sub-process p_j, this approach estimates the errors for a false alarm, α_j, and non-detection, β_j, based on the error sources within the respective sub-processes. Assuming independence of error probabilities among the sub-processes, the overall errors can then be calculated as

$$\alpha_{total} = 1 - \prod_{j=1}^{5}(1 - \alpha_j) \qquad \beta_{total} = 1 - \prod_{j=1}^{5}(1 - \beta_j)$$

[4] *Intelligence Cycle*. Retrieved from https://www.fbi.gov/about-us/intelligence/intelligence-cycle.