Statistical framework for decision making in mine action

Jan Larsen
Intelligent Signal Processing Group
Informatics and Mathematical Modelling
Technical University of Denmark
jl@imm.dtu.dk, www.imm.dtu.dk/~jl
Why do we need statistical models and machine learning?

- Mine action is influenced by many uncertain factors
- The goals of mine action depends on difficult socio-economic and political considerations

Scientist are born sceptical: they don’t believe facts unless they see them often enough
Why do we need statistical models and machine learning?

- Statistical modeling is the **principled framework** to handle uncertainty and complexity.
- Statistic modeling usually focuses on identifying important parameters.
- Machine learning learns complex models from collections of data to make optimal predictions in new situations.
Why do we need statistical models and machine learning?

- Statistical modeling is the **principled framework** to handle uncertainty and complexity.
- Statistical modeling usually focuses on identifying important parameters.
- Machine learning learns complex models from collections of data to make optimal predictions in new situations.
- Facts, prior information, consistent and robust information and decisions with associated risk estimates.
There is no such thing as facts to spoil a good explanation!

- Pitfalls and misuse of statistical methods sometimes wrongly leads to the conclusion that they are of little practical use.

After the dogs went in we never saw an accident.

Most suspected areas have very few mines.
There is no such thing as facts to spoil a good explanation!

- Pitfalls and misuse of statistical methods sometimes wrongly leads to the conclusion that they are of little practical use.
- Some data are in the tail of the distribution: generalization from few examples is not possible.
- Smoking is not dangerous: my granny just turned 95 and has been a heavy smoker all his live.
The elements of statistical decision theory

**Data**
- Sensor
- Calibration
- Post clearance
- External factors

**Prior knowledge**
- Physical knowledge
- Experience
- Environment

**Statistical models**

**Loss function**
- Decisions
- Risk assessment

**Inference:** assign probabilities to hypotheses about the suspected area
Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction

- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse project
Detecting a mine – tossing a coin

\[ Frequency = \frac{\text{no of heads}}{\text{no of tosses}} \]

\[ \text{probability} = \text{frequency} \text{ when infinitely many tosses} \]
On 99.6% detection probability

\[ Frequency = \frac{996}{1000} = 99.6\% \]

One more (one less) count will change the frequency a lot!

\[ Frequency = \frac{9960}{10000} = 99.60\% \]
Detection probability - tossing a coin

- $N$ independent tosses number of
- $y$ number of heads observed
- $\theta$ probability of heads

$$\hat{\theta} = \frac{y}{N}$$

$$P(y \mid \theta) = \text{Binom}(\theta \mid N) = \binom{N}{y} \theta^y \theta^{N-y}$$

Data likelihood
Prior beliefs and opinions

- Prior 1: the fair coin: $\theta$ should be close to 0.5
- Prior 2: all values of $\theta$ are equally plausible

$$p(\theta) = \text{Beta}(\theta | \alpha, \beta)$$
Prior beliefs and opinions

\[ p(\theta) = \begin{cases} 
1, & \alpha = 1, \beta = 1 \\
1, & \alpha = 3, \beta = 3 
\end{cases} \]
Bayes rule: combining data likelihood and prior

\[ P(\theta \mid y) = \frac{P(y \mid \theta)p(\theta)}{P(y)} \]

\[ P(\theta \mid y) = \text{Beta}(\theta \mid y + \alpha, \beta + n - y) \sim \theta^{y+\alpha} \theta^{n-y+\beta} \]
Posterior probability is also Beta

\[ P(\theta \mid y) = \text{Beta}(\theta \mid y + \alpha, \beta + n - y) \sim \theta^{y+\alpha} \theta^{n-y+\beta} \]
Posterior after observing one head

Flat prior

\[ \text{Beta}(\theta | 2, 1) \]

mean = 2/3

Fair coin

\[ \text{Beta}(\theta | 4, 3) \]

mean = 4/7
Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction

- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse project
What are the requirements for mine action risk

- Tolerable risk for individuals comparable to other natural risks
- As high cost efficiency as possible requires detailed risk analysis – e.g. some areas might better be fenced than cleared
- Need for professional risk analysis, communication management and control involving all partners (MAC, NGOs, commercial etc.)
What are the requirements for mine action risk

- Tolerable risk for individuals comparable to other natural risks

**Fact**

99.6% is not an unrealistic requirement but... today’s methods achieve at most 90% and are hard to evaluate!!!

GICHD and FFI are currently working on such methods [Håvard Bach, Ove Dullum NDRF SC2006]
A simple inference model – assigning probabilities to data

- The detection system provides the probability of detection a mine in a specific area: \( \text{Prob(detect)} \)
- The land area usage behavior pattern provides the probability of encounter: \( \text{Prob(mine encounter)} \)

\[
\text{Prob(casualty)} = (1 - \text{Prob(detect)}) \times \text{Prob(mine encounter)}
\]

For discussion of assumptions and involved factors see
“Risk Assessment of Minefields in HMA – a Bayesian Approach”

A simple loss/risk model

- Minimize the number of casualties
- Under mild assumptions this equivalent to minimizing the probability of casualty
Requirements on detection probability

\[
\text{Prob(causeality)} = (1 - \text{Prob(detection)}) \times \text{Prob(encounter)}
\]

\[
\text{Prob(detection)} = 1 - \frac{\text{Prob(causeality)}}{\text{Prob(encounter)}}
\]

- \text{Prob(encounter)} = \rho \times a
  - \rho: \text{homogeneous mine density (mines/m}^2\text{), a: yearly footprint area (m}^2\text{)}
- \text{Prob(causeality)} = 10^{-5} \text{ per year}
### Maximum yearly footprint area in m²

<table>
<thead>
<tr>
<th>P(detection)</th>
<th>(\rho) : mine density (mines/km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.996</td>
<td>25000   2500   250   25   2.5</td>
</tr>
<tr>
<td>0.9</td>
<td>1000    100    10    1    0.1</td>
</tr>
</tbody>
</table>

**Reference:** Bjarne Haugstad, FFI
Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction

- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse project
Evaluation and testing in MA

- How do we assess the performance/detection probability?
- What is the confidence?

Overfitting
- insufficient coverage of data
- unmodeled confounding factors
- insufficient model fusion and selection

Changing environment
- mine types, placement
- soil and physical properties
- unmodeled confounds
Two types of error in detection of mines

**Sensing error**

The system does not sense the presence of the mine object

- decrease in detection probability

**Decision error**

The detector misinterprets the sensed signal

- increase in false alarm rate
Two types of error in detection of mines

Sensing error
- The system does not sense the presence of the mine object
- Sensing error: The mine has low metal content
- Decision error: The detector misinterprets the sensed signal

Example: metal detector
- Sensing error: the mine has low metal content
- Decision error: a piece of scrap metal was found

Example: mine detection dog
- Sensing error: the TNT leakage from the mine was too low
- Decision error: the dog handler misinterpreted the dogs indication
Confusion matrix in system design and test phase which should lead to certification

<table>
<thead>
<tr>
<th>Estimated</th>
<th>True</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>a</td>
</tr>
<tr>
<td>no</td>
<td>c</td>
</tr>
</tbody>
</table>

- Detection probability (sensitivity): $\frac{a}{a+c}$
- False alarm: $\frac{b}{a+b}$
- False positive (specificity): $\frac{b}{b+d}$
Receiver operation characteristic (ROC)

detection probability %

false alarm %

0 100

0 100
Inferring the detection probability

- $N$ independent mine areas for evaluation
- $y$ detections observed
- True detection probability $\theta$

$$P(y \mid \theta) \sim \text{Binom}(\theta \mid N) = \binom{N}{y} \theta^{y} \theta^{N-y}$$
Bayes rule: combining data likelihood and prior

\[ P(\theta \mid y) = \frac{P(y \mid \theta)p(\theta)}{P(y)} \]

\[ P(\theta \mid y) = \text{Beta}(\theta \mid y + \alpha, \beta + n - y) \sim \theta^{y+\alpha} \theta^{n-y+\beta} \]
Prior distribution

\[ p(\theta) \]

- Green line: \( \alpha = 1, \beta = 1 \)
- Red line: \( \alpha = 0.9, \beta = 0.6 \)

Mean = 0.6
HPD credible sets – the Bayesian confidence interval

\[ C_{1-\varepsilon} = \{ \theta : \text{P}(\theta | y) \geq k(\varepsilon) \} \], \quad \text{CDF}(\theta | y) > 1 - \varepsilon

\[ \alpha \approx 32.9, \beta \approx 18.6 \]

N=50, y=32, \( \theta_{est} = 0.64 \)

\( C_{95} = 0.52665, C_{99} = 0.47862 \)
The required number of samples $N$

- We need to be confident about the estimated detection probability

$$\text{Prob}(\theta > 99.6\%) = C_{1-\epsilon}$$

<table>
<thead>
<tr>
<th>$\theta_{\text{est}}$</th>
<th>$C_{95%}$</th>
<th>$C_{99%}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>99.7%</td>
<td>9303</td>
<td>18994</td>
</tr>
<tr>
<td>99.8%</td>
<td>2285</td>
<td>3995</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\theta_{\text{est}}$</th>
<th>$C_{95%}$</th>
<th>$C_{99%}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>99.7%</td>
<td>8317</td>
<td>18301</td>
</tr>
<tr>
<td>99.8%</td>
<td>2147</td>
<td>3493</td>
</tr>
</tbody>
</table>

Uniform prior

Informative prior

$\alpha = 0.9, \beta = 0.6$
Credible sets when detecting 100%

Minimum number of samples $N$

<table>
<thead>
<tr>
<th></th>
<th>Prob($\theta &gt; 80%$)</th>
<th>Prob($\theta &gt; 99.6%$)</th>
<th>Prob($\theta &gt; 99.9%$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{95%}$</td>
<td>13</td>
<td>747</td>
<td>2994</td>
</tr>
<tr>
<td>$C_{99%}$</td>
<td>20</td>
<td>1148</td>
<td>4602</td>
</tr>
</tbody>
</table>
Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction

- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse project
Efficient MA by hierarchical approaches

Ref: Håvard Bach, Paul Mackintosh
Danger maps

- The outcome of a hierarchical surveys
- Information about mine types, deployment patterns etc. should also be used
- Could be formulated/interpreted as a prior probability of mines

SMART system described in GICHD: Guidebook on Detection Technologies and Systems for Humanitarian Demining, 2006
Sequential information gathering

technical survey

mine clearance

data

data

prior

posterior

prior

posterior
Statistical information aggregation

- $e=1$ indicates encounter of a mine in a box at a specific location
- probability of encounter $P(e = 1)$ from current danger map
- $d=1$ indicates detection by the detection system
- probability of detection $P(d = 1)$ from current accreditation

\[
P(e = 1 \land d = 0) = P(e = 1)(1 - P(d = 1))
\]

$P(\text{no mine}) = 1 - P(e = 1 \land d = 0)$
Statistical information aggregation

Example: flail in a low danger area

\[ P(e = 1) = 0.2, \ P(d = 1) = 0.8 \]
\[ P(\text{no mine}) = 1 - P(e = 1 \land d = 0) = 1 - 0.2 \times 0.2 = 0.96 \]

Example: manual raking in a high danger area

\[ P(e = 1) = 1, \ P(d = 1) = 0.96 \]
\[ P(\text{no mine}) = 1 - P(e = 1 \land d = 0) = 1 - 1 \times 0.04 = 0.96 \]
Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction

- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse project
Improving performance by fusion of methods

- Methods (sensors, mechanical etc.) supplement each other by exploiting different aspects of the physical environment

  - Late integration
  - Hierarchical integration
  - Early integration
Early integration – sensor fusion

Sensor 1 → Trainable sensor fusion → Detection
Sensor n → Trainable sensor fusion

database
Late integration – decision fusion

Sensor → Signal processing → Decision fusion

Mechanical system → Decision
Advantages

- Combination leads to a possible exponential increase in detection performance
- Combination leads to better robustness against changes in environmental conditions
Challenges

- Need for certification procedure of equipment under well-specified conditions (ala ISO)
- Need for new procedures which estimate statistical dependences between existing methods
- Need for new procedures for statistically optimal combination
Outline

The design and evaluation of mine clearance equipment – the problem of reliability
- Detection probability – tossing a coin
- Requirements in mine action
- Detection probability and confidence in MA
- Using statistics in area reduction

Improving performance by information fusion and combination of methods
- Advantages
- Methodology
- DeFuse project
### Dependencies between methods

#### Contingency tables

<table>
<thead>
<tr>
<th>Mine present</th>
<th>Method j</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>no</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method i</th>
<th>Method j</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>c11</td>
</tr>
<tr>
<td>no</td>
<td>c01</td>
</tr>
</tbody>
</table>
Optimal combination

Method 1 → 0/1 → Combiner → 0/1
Method K → 0/1

Optimal combiner depends on contingency tables
**Optimal combiner**

<table>
<thead>
<tr>
<th>Method</th>
<th>Combiner</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>

OR rule is optimal for independent methods

\[ 2^{2^{K-1}} - 1 \] possible combiners
OR rule is optimal for independent methods

<table>
<thead>
<tr>
<th>Method 1</th>
<th>1 0 0 1 0 0 1 0 1 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method 2</td>
<td>0 1 0 0 1 0 1 1 1 0</td>
</tr>
<tr>
<td>Combined</td>
<td>1 1 0 1 1 0 1 1 1 0</td>
</tr>
</tbody>
</table>

\[
P_d(\text{OR}) = P(\hat{y}_1 \lor \hat{y}_2 = 1 \mid y = 1) \\
= 1 - P(\hat{y}_1 = 0 \land \hat{y}_2 = 0 \mid y = 1) \\
= 1 - P(\hat{y}_1 = 0 \mid y = 1) \cdot P(\hat{y}_2 = 0 \mid y = 1) \\
= 1 - (1 - P_{d1}) \cdot (1 - P_{d2})
\]
False alarm follows a similar rule

\[
P_{fa}(OR) =
\]
\[
P(\hat{y}_1 \lor \hat{y}_2 = 1 \mid y = 0)
\]
\[
= 1 - P(\hat{y}_1 = 0 \land \hat{y}_2 = 0 \mid y = 0)
\]
\[
= 1 - P(\hat{y}_1 = 0 \mid y = 0) \cdot P(\hat{y}_2 = 0 \mid y = 0)
\]
\[
= 1 - (1 - P_{fa1}) \cdot (1 - P_{fa2})
\]
Example

\[ p_{d1} = 0.8, p_{fa1} = 0.1 \quad p_{d2} = 0.7, p_{fa2} = 0.1 \]

\[ p_d = 1 - (1 - 0.8) \cdot (1 - 0.7) = 0.94 \]

\[ p_{fa} = 1 - (1 - 0.1) \cdot (1 - 0.1) = 0.19 \]

Exponential increase in detection rate
Linear increase in false alarm rate

Joint discussions with: Bjarne Haugstad
Testing independence – Fisher’s exact test

<table>
<thead>
<tr>
<th>Method i</th>
<th>Method j</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>no</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>no</td>
</tr>
</tbody>
</table>

**Hypothesis:** Method i and j are independent

**Alternatives:** Dependent or positively (negatively) correlated

\[
H : P(\hat{y}_i = 0, \hat{y}_j = 0) = P(\hat{y}_i = 0) \cdot P(\hat{y}_j = 0)
\]

\[
A : P(\hat{y}_i = 0, \hat{y}_j = 0) > P(\hat{y}_i = 0) \cdot P(\hat{y}_j = 0)
\]
Artificial example

- N=23 mines
- Method 1: P(detection)=0.8, P(false alarm)=0.1
- Method 2: P(detection)=0.7, P(false alarm)=0.1
- Resolution: 64 cells

How does detection and false alarm rate influence the possibility of gaining by combining methods?
Confusion matrix for method 1

<table>
<thead>
<tr>
<th></th>
<th>True</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>Estimated</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>no</td>
</tr>
</tbody>
</table>
Confidence of estimated detection rate

With $N=23$ mines 95%-credible intervals for detection rates are extremely large!!!!

Method 1 (flail): [64.5% 82.6% 93.8%]

Method 2 (MD): [50.4% 69.6% 84.8%]
Confidence for false alarm rates

- Determined by deployed resolution
- Large resolution - many cells gives many possibilities to evaluate false alarm.
- In present case: 64-23=41 non-mine cells

Method1 (flail): [4.9% 12.2% 24.0%]
Detection rates

Flail: 82.6
Metal detector: 69.6
Combined: 91.3
False alarm rates

Flail: 12.2  
Metal detector: 7.3  
Combined: 17.1

Combination number

%
Comparing methods

- Is the combined method better than any of the two original?
- Since methods are evaluated on same data a paired statistical McNemar with improved power is useful

Method 1 (flail): 82.6% < 91.3% Combined

Method 2 (MD): 69.6% < 91.3% Combined
Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction

- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse project
They keys to a successful mine clearance system

- Use statistical learning which combines all available information in an optimal way
  - informal knowledge
  - data from design test phase
  - confounding parameters (environment, target, operational)

- Combine many different methods using statistical fusion

MineHunt System and HOSA concepts have been presented at NDRF summer conferences (98,99,01)
Obtain general scientific knowledge about the advantages of deploying a combined approach

Eliminate confounding factors through careful experimental design and specific scientific hypotheses

Test the general scientific hypothesis is that there is little dependence between missed detections in successive runs of the same or different methods

To accept the hypothesis under varying detection/clearance probability levels

To lay the foundation for new practices for mine action, but it is not within scope of the pilot project

Systems: ALIS dual sensor, MD, MDD, Hydrema flail
Conclusions

- Statistical decision theory and modeling is essential for optimal use of prior information and empirical evidence.
- It is very hard to assess the necessary high performance which is required to have a tolerable risk of casualty.
- The use of sequential information aggregation is promising for developing new hierarchical survey schemes (SOPs).
- Combination of methods is a promising avenue to overcome current problems.

Jan Larsen 66