

# Meerarmige bandieten

**Begeleider: Lodewijk Kallenberg**

Beschouw de volgende situatie: er is een aantal projecten, elk project bevindt zich in een bepaalde toestand, en aan het begin van iedere dag moet worden bepaald aan welk project die dag wordt gewerkt. De toestand van het project waaraan wordt gewerkt verandert (in het algemeen stochastisch), terwijl de toestand van de projecten waaraan niet wordt gewerkt onveranderd blijft. Bovendien ontvangt men een opbrengst die alleen afhankelijk is van het project waaraan wordt gewerkt, en dus niet van de overige projecten. De naam *bandiet* komt voort uit de analogie met 'eenarmige bandieten' ofwel 'fruitmachines'.

*De vraagstelling luidt: gegeven de toestanden van de diverse projecten, de veranderings- en de opbrengststructuur, aan welk project kan het beste worden gewerkt om de zogenaamde verdisconteerde opbrengsten, d.w.z. de opbrengsten waarbij rente in het spel is, te maximaliseren?*

Er zijn allerlei toepassingen van 'meerarmige bandieten': in de farmacie (waarbij de projecten verschillende medicijnen zijn om een kwaal mee te bestrijden), in de scheduling (waar de projecten machines zijn) en bij het zoeken naar goud (waar de projecten mogelijke vindplaatsen zijn).

Het 'meerarmige bandiet' probleem kan worden gemodelleerd als een Markov beslissingsprobleem. Formeel is dit model als volgt:

- $S$  is een eindige *toestandsruimte*, zeg  $S = \{1, 2, \dots, N\}$ ;
- In toestand  $i \in S$  wordt een actie uit een eindige *actieverzameling*  $A(i)$  gekozen;
- Indien in toestand  $i$  actie  $a \in A(i)$  wordt gekozen, dan gebeurt het volgende:
  - (1) er is een *directe opbrengst*  $r_i(a)$ ;
  - (2) met *overgangskans*  $p_{ij}(a)$  is het systeem op het volgende beslissingstijdstip in toestand  $j$ .

Het probleem zit vooral in de complexiteit van de toestandsruimte: bij 10 projecten die elk 10 mogelijke toestanden hebben heeft de totale toestandsruimte  $10^{10}$  elementen. Gelukkig is het mogelijk een oplossing te vinden door de diverse projecten afzonderlijk te beschouwen, wat in het zojuist genoemde voorbeeld betekent dat 10 keer een probleem met 10 toestanden moet worden opgelost, wat het vinden van een overall oplossing veel eenvoudiger maakt.

Het project houdt het volgende in:

1. Het een en ander bestuderen over Markovbeslissingsproblemen.
2. Een aantal artikelen bestuderen waarin oplossingsmethoden worden voorgesteld, onder andere een recent artikel uit 2007.
3. Een vergelijkend experimenteel onderzoek doen naar de diverse methoden.