

Bayesian Network

Joanna Kołodziejczyk

Marzec 2021

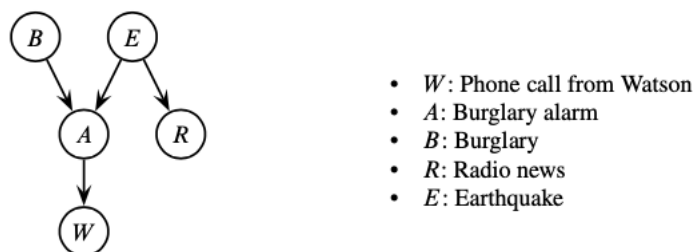
1 Scenariusz

Pan Holmes mieszka na obszarze o wysokiej przestępczości i dlatego zainstalował alarm przeciw włamaniowy. Holmes ufa, że jego sąsiedzi zadzwonią do niego, gdy usłyszą sygnał alarmu. Pan Holmes ma sąsiada dr Watsona.

Niestety, sąsiad nie do końca jest wiarygodny. Dr Watson jest znany jako żartowniś.

Pan Holmes wie również z instrukcji obsługi swojego systemu alarmowego, że urządzenie jest wrażliwe na trzęsienia ziemi i może zostać wyzwolone przez nie przypadkowo. Zdaje on sobie sprawę, że gdyby doszło do trzęsienia ziemi, na pewno byłoby ono w wiadomościach radiowych.

Pan Holmes pracuje w swoim biurze, kiedy odbiera telefon od swojego sąsiada, doktora Watsona, który mówi mu, że włączył się alarm przeciw włamaniowy Holmesa. Przekonany, że włamywacz włamał się do jego domu, Holmes spieszy się do samochodu i jedzie do domu. Po drodze słucha radia, a w wiadomościach donosi się, że w okolicy doszło do niewielkiego trzęsienia ziemi. Wiedząc, że trzęsienia ziemi mają tendencję do uruchamiania alarmów przeciw włamaniowych, wraca do swojej pracy.



Wszystkie zmienne są zmiennymi dyskretnymi binarnymi. Prawdopodobieństwo a priori węzłów B i E, to:

Tablica 1: Burglary prawdopodobieństwo a priori

yes	no
0.3	0.7

Tablica 2: Earthquake prawdopodobieństwo a priori

yes	no
0.35	0.65

Wszystkie węzły, które mają rodziców mają podane rozkłady prawdopodobieństw warunkowych (a posteriori) węzłów A, R i W, to:

Tablica 3: Alarm tablica rozkładów warunkowych

Burglary yes				Burglary no			
Earthquake yes		Earthquake no		Earthquake yes		Earthquake no	
Alarm y	Alarm n	Alarm y	Alarm n	Alarm y	Alarm n	Alarm y	Alarm n
0.95	0.05	0.9	0.1	0.6	0.4	0.01	0.99

Tablica 4: News tablica rozkładów warunkowych

Earthquake yes		Earthquake no	
News y	News no	News y	News n
0.6	0.4	0.01	0.99

Tablica 5: Watson tablica rozkładów warunkowych

Alarm yes		Alarm no	
Watson y	Watson n	Watson y	Watson n
0.8	0.2	0.4	0.6

Przekazana struktura i prawdopodobieństwa w każdym węźle opisują wiedzę o problemie i pozwalają na probabilistyczne wnioskowanie.

2 Implementacja sieci „Alarm” w bibliotece bnlearn

Poniższy kod pozwala przedstawić sieć Alarm z użyciem poleceń biblioteki bnlearn.

Instrukcja instalacji biblioteki wraz z jej niezbędnymi elementami jest podana na stronie domowej biblioteki <http://www.bnlearn.com>.

2.1 Tworzenie struktury sieci Bayesowskiej

Na początek tworzymy strukturę acyklicznego grafu skierowanego repetującego wiedzę o zależności zdarzeń w wybranym problemie. Nieśniejszy przykład dotyczy analizy przypadku p. Holmesa.

```

1 library(bnlearn)
2 library(Rgraphviz)
3
4 #Formuła na tworzenie sieci połączeń poprzez wskazanie węzłów
5 #i ich powiązań warunkowych
6 bl.alarm <- model2network('[Burglar][Earthquake][Alarm|Burglar:
7   Earthquake]
8   [News|Earthquake][Watson|Alarm]')
9 plot(bl.alarm)
10
11 #Alternatywna forma definiowania grafu poprzez wytworzenie powiązań
12   metoda set.arc
13 bl.dag = empty.graph(nodes = c("B", "E", "N", "A", "W"))
14 bl.dag = set.arc(bl.dag, from = "E", to = "N")
15 bl.dag = set.arc(bl.dag, from = "E", to = "A")
16 bl.dag = set.arc(bl.dag, from = "B", to = "A")
17 bl.dag = set.arc(bl.dag, from = "A", to = "W")
18 bl.dag = set.arc(bl.dag, from = "E", to = "N")
19 plot(bl.dag)
20
21 #Alternatywna forma definiowania grafu poprzez połączenie
22 #wszystkich węzłów za pomocą macierzy powiązań
23 survey.dag = empty.graph(nodes = c("Bu", "Er", "Ne", "Al", "Wa"))
24 arc.set = matrix(c("Bu", "Al",
25                   "Er", "Al",
26                   "Er", "Ne",
27                   "Al", "Wa"),
28                 byrow = TRUE, ncol = 2,
29                 dimnames = list(NULL, c("from", "to")))
30 arcs(survey.dag) = arc.set
31 plot(survey.dag)

```

2.2 Modyfikowania struktury sieci

Na wytworzonej lub wczytanej raz strukturze można dokonywać modyfikacji, dokonywać różnych kontroli i ją przedstawić graficznie. Biblioteka dba, by nie utworzyć grafu cyklicznego. Poniższy przykład pokazuje sposoby manipulowania połączeniami w grafie i kontrolę lokalnych zależności. Można sprawdzać, czy w grafie są ścieżki pomiędzy wskazanymi węzłami.

```

1 #usuwanie krawędzi w grafie
2 modified.alarm <- drop.arc(bl.alarm, "Earthquake", "News")
3 plot(modified.alarm)
4 #zmiana kierunku krawędzi w grafie
5 modified.alarm <- reverse.arc(modified.alarm, "Alarm", "Burglar")
6 plot(modified.alarm)
7 #tworzenie nowej krawędzi w grafie
8 modified.alarm <- set.arc(modified.alarm, "Earthquake", "Burglar")
9 plot(modified.alarm)
10
11 modified.alarm <- set.arc(modified.alarm, "Watson", "Earthquake")
12 # Acykliczność jest domyślnie wymuszona przez wszystkie funkcje
13   tworzące graf
14 # W konsekwencji powstanie błąd i

```

```

14 #Error in arc.operations(x = x, from = from, to = to, op = "set",
15 #check.cycles = check.cycles, : the resulting graph contains
    cycles.
16
17
18 bl.alarm
19 # Wyświetlenie sąsiadów węzła
20 nbr(bl.alarm, node = 'Alarm')
21 # Wyświetlenie rodziców węzła
22 parents(bl.alarm, node = 'Watson')
23 # Wyświetlenie potomków węzła
24 children(bl.alarm, node = 'Alarm')
25 # Wyświetlenie „Markov blanket”
26 mb(bl.alarm, node = 'Alarm')
27
28 # Możemy dostosować formę wizualizacji
29 plot(bl.alarm, highlight = list(nodes='Alarm'))
30 # z użyciem biblioteczki Rgraphviz
31 hlight <- list(nodes = c("Earthquake"), arcs = c("Earthquake", "
    News"),
32   col = "blue", textCol = "grey")
33 pp <- graphviz.plot(bl.alarm, highlight = hlight)
34 Rgraphviz::renderGraph(pp)
35
36
37 #Możemy wyświetlić wszystkie luki lub poprosić o ścieżkę między
    para węzłów:
38 arcs(bl.alarm)
39
40 path(bl.alarm, from = "Burglar", to = "Watson")
41 path(bl.alarm, from = "Watson", to = "Burglar")

```

2.3 Podanie wartości rozkładów w węzłach

Przekazanie listy lokalnych rozkładów warunkowych do `custom.fit()` generuje `BayesNet.fit` (w przykładzie `bl.alarm.fit`). Dla sieci ze zmiennymi dyskretnymi możemy określić CPT (Conditional Probability Table) dla każdego węzła: Kolejność podanych prawdopodobieństw jest zgodna z kolejnością podaną w tabelach w poprzedniej sekcji.

```

1 # Zmienna yn przechowuje wartości dla wszystkich binarnych
    zmiennych losowych
2 yn <- c("yes", "no")
3 # Rozkład dla węzła Burglary
4 B <- array(dimnames = list(Burglar = yn), dim = 2, c(0.30,0.70))
5 # Rozkład dla węzła Earthquake
6 E <- array(dimnames = list(Earthquake = yn), dim = 2, c(0.35,0.65))
7 # Rozkład dla węzła Alarm
8 A <- array(dimnames = list(Alarm = yn, Earthquake = yn, Burglar =
    yn),
9   dim = c(2, 2, 2),
10  c(0.95,0.05,0.90,0.10,0.60,0.40,0.01,0.99))
11 W <- array(dimnames = list(Watson = yn, Alarm = yn),
12  dim = c(2, 2), c(0.80,0.20,0.40,0.60))
13 N <- array(dimnames = list(News = yn, Earthquake = yn),

```

```

14         dim = c(2, 2), c(0.60,0.40,0.01,0.99))
15 cpts <- list(Burglar = B, Earthquake = E, Alarm = A, Watson = W,
16             News = N)
17 bl.alarm.fit = custom.fit(bl.alarm, cpts)
18 # Co przechowuje nauczony model Sieci Bayesa
19 bl.alarm.fit
20 # Co przechowuje nauczony model Sieci Bayesa dla zmiennej
21   Earthquake
22 bl.alarm.fit$Earthquake
23 #Wykresy prawdopodobienstw dla poszczegolnych wezlow
24 bn.fit.barchart(bl.alarm.fit$Earthquake)
25 bn.fit.barchart(bl.alarm.fit$News)
26 bn.fit.barchart(bl.alarm.fit$Alarm)

```

2.4 Wnioskowanie

Użyjemy pakietu gRain do wnioskowania dokładnego (exact inference). Możemy przekonwertować obiekty bnlearn na gRain. Czy potrzebujemy bl.alarm lub bl.alarm.fit do wnioskowania?

```

1 # Wnioskowanie pelne
2 gr.alarm <- as.grain(bl.alarm.fit)
3 # Aby obliczyc wszystkie prawdopodobienstwa w sieci nalezy wykona
4   kompilacje
5 gr.alarm <- compile(gr.alarm)
6 gr.alarm
7 #Mozna teraz pytac o prawdopodobienstwa kazdego wezla w sieci
8 querygrain(object=gr.alarm, nodes="Earthquake")
9 querygrain(object=gr.alarm, nodes="Watson")
10 querygrain(object=gr.alarm, nodes="Alarm")
11 querygrain(object=gr.alarm, nodes=c('News','Watson'))

```

Użyjemy setEvidence() i pEvidence(), querygrain(), aby uzyskać prawdopodobieństwo, gdy przedstawiono dowody.

```

1 #Wnioskowanie, gdy wszystkie zdarzenia maja wartosc 'no'
2 no <- rep("no", 5)
3 nodes <- c('Burglar', 'Earthquake', 'Alarm', 'Watson', 'News')
4 gr.alarm <- setEvidence(object=gr.alarm, nodes=nodes, states=no)
5 gr.alarm
6 pEvidence(gr.alarm) # Prawdopodobienstwo zdarzenia to
7
8 #Powrot do ustawien (bez dowodow)
9 gr.alarm <- retractEvidence(gr.alarm, nodes)
10 gr.alarm
11
12 # Pytanie jakie jest prawdopodobienstwo wlamania, gdy Watson dzwoni
13   ?
14 # 1. setEvidence
15 # 2. querygrain
16 gr.alarm.Wats <- setEvidence(object=gr.alarm, nodes = "Watson",
17   states = "yes")
18 querygrain(gr.alarm.Wats, nodes = c('Burglar'), type = "joint")

```

3 Zadania

Należy na podstawie podanych fragmentów kodu wykonać obliczenia:

1. Zadania do problemu pana Holmes'a.
 - (a) Dowód: Watson zadzwonił. Odpowiedzieć na pytania:
 - i. Czy prawdopodobieństwo Earthquake jest teraz większe? $P(E|W) > P(E)$?
 - ii. Czy prawdopodobieństwo Włamania jest teraz większe? $P(B|W) > P(B)$?
 - (b) Dowód: Jest trzęsienie ziemi (Earthquake = yes). Odpowiedzieć na pytania:
 - i. Czy prawdopodobieństwo, że Watson zadzwoni jest teraz większe? $P(W|E) > P(W)$?
 - ii. Czy prawdopodobieństwo Włamania jest teraz większe? $P(B|E) > P(B)$? Dlaczego tak lub dlaczego nie?
 - (c) Dowód: Alarm dzwoni (Alarm = yes). Odpowiedzieć na pytania:
 - i. Czy prawdopodobieństwo Włamania jest teraz większe? $P(B|A) > P(B)$?
 - (d) Dowód: Alarm dzwoni i jest trzęsienie ziemi (Alarm = yes, Earthquake = yes). Odpowiedzieć na pytania:
 - i. Czy prawdopodobieństwo Włamania jest teraz większe? $P(B|A) < P(B|A, E)$?

UWAGA: Przykład deklarowania węzła z dyskretną zmienną losową z 3 wartościami:

```
1 A.lv = c("young", "adult", "old")
2 A.prob = array(c(0.30, 0.50, 0.20), dim = 3, dimnames = list(A = A.lv))
```

4 Rozliczenie zadania

Jako rozliczenie do każdego zadania należy podać:

1. fragment kodu tworzący założenie (evidence) oraz wnioski (pytanie o wartość zmiennej losowej)
2. odpowiedź na zadanie pytanie oraz uzasadnienie i interpretację takiej właśnie zmiany.

Bardzo proszę o zamieszczenie sprawozdania w Teams. Nazwa pliku powinna zawierać imię i nazwisko.