Razširjanje zaupanja za lokalizacijo senzorskih omrežji (osnutek)

Jaka Velkaverh
Mentorja: Sergio Cabello, Tomaž Javornik

28. februar 2024

Povzetek

V osnutku opišemo teoretično osnovo za postopek lokalizacije senzorskih omrežji z metodo razširjanja zaupanja. Formuliramo pojme potrebne za izpeljavo postopka in ga izpeljemo v primeru, da je omrežje drevo.

1 Uvod

1.1 Cilj

V celem delu bomo predstavili algoritem za lokalizacijo senzorskih omrežji. Senzorsko omrežje je sestavljeno iz več naprav. Napravam, katerih pozicija je znana, pravimo sidra, ostalim pa agenti. Naprave izmerijo približno medsebojno razdaljo, a je nekateri pari naprav zaradi prevelike razdalje ali pregrad ne morejo. Iz točnih leg sider in izmerjenih razdalj želimo izračunali približke za lege agentov.

1.2 Oznake

Omrežju naprav pripišemo graf $G = (V, E)$, kjer so vozlišča naprave. Napravi sta povezani, če je bila njuna medsebojna razdalja izmerjena. Lokacijo naprav $v \in V$ obravnavamo kot slučajno spremenljivko X_v. To je dvorazsežni ali trirazsežni slučajni vektor, a ga zaradi preglednosti ne bomo podčrtovali. Za neko podmnožico vozlišč $S = \{v_1, \ldots, v_k\} \subseteq V$ označimo z $\underline{X}_S$ slučajni vektor $(X_{v_1}, \ldots, X_{v_k})$.

2 Markovska polja

Osnovna lastnost, ki bo vodila v Markovska polja je pogojna neodvisnost. Pogojna neodvisnost je šibkejša lastnost od neodvisnosti.

Naj bodo X, Y, Z slučajne spremenljivke. Z pogojno neodvisnostjo X in Y ob Z hočemo povedati, da X in Y nimata neposrednega vpliva drug na drugega in da je edini posredni vpliv tisti, ki gre „skozi“ Z. Torej če fiksiramo vrednost Z, bosta spremenljivki X in Y neodvisni. Vrednosti ene nam nič ne povejo o vrednostih druge, saj je posreden vpliv „blokiran“ z fiksiranim Z.

Definicija 2.1. *Nakjučna vektorja $\underline{X}, \underline{Y}$ s pogojnima gostotama $f_{\underline{X}|\underline{Z}}$, $f_{\underline{Y}|\underline{Z}}$ sta pogojno neodvisna ob $\underline{Z}$, če za skupno pogojno gostoto velja:*

$$\forall \underline{z}. \quad f_{\underline{X},\underline{Y}|\underline{Z}}\left(\underline{x}, \underline{y} \mid \underline{Z} = \underline{z}\right) = f_{\underline{X}|\underline{Z}}\left(\underline{x} \mid \underline{Z} = \underline{z}\right) \; f_{\underline{Y}|\underline{Z}}\left(\underline{y} \mid \underline{Z} = \underline{z}\right)$$

Z tem pojmom lahko definiramo Markovsko polje. Glavna ideja je, da želimo z grafom ponazoriti medsebojne odvisnosti več naključnih spremenljivk. Ne moremo zahtevati, da so nesosednje spremenljivke neodvisne, ker so lahko spremenljivke iz iste komponente za povezanost odvisne posredno. Če pa fiksiramo vse ostale spremenljivke, pa lahko pričakujemo, da bosta nesosednji spremenljivki neodvisni, saj smo se znebili posrednih vplivov.

Definicija 2.2. *Naj bo $G = (V, E)$, $V = \{v_1, v_2, \ldots, v_n\}$ in slučajne spremenljivke $\underline{X}_V = (X_{v_1}, X_{v_2}, \ldots, X_{v_n})$. Spremenljivke $\underline{X}_V$ in graf G tvorijo* Markovsko polje*, ko sta za vsaka nesosedna $u, v \in V$, $uv \notin E$ pripadajoči spremenljivki X_u in X_v pogojno neodvisni ob vseh ostalih spremenljivkah $\underline{X}_{V\setminus\{u,v\}}$.*

Opomba. *Pogoj zgoraj se imenuje parna lastnost Markova. Obstajata še lokalna in globalna lastnost Markova, ki sta strožji, a ju v tem delu ne potrebujemo.*

Izrek 2.3 (Hammersley–Clifford)**.** *Naj bo $G = (V, E)$ graf in $\underline{X}_V$ zvezno porazdeljene naključne spremenljivke, ki pripadajo vozliščem. Dodatno $\forall \underline{x}_V.\ f_{\underline{X}_V}(\underline{x}_V) > 0$. Tedaj G in $\underline{X}_V$ tvorita Markovsko polje natanko tedaj, ko se skupna gostota faktorizira po klikah grafa G:*

$$f_{\underline{X}_V}(\underline{x}_V) \propto \prod_{C \text{ klika } G} \psi_C(\underline{x}_C)$$

Ni dokaza. Nekaj komentarjev glede izreka:

- $f(x) \propto g(x)$ pomeni $\exists C \in \mathbb{R}\setminus\{0\}.\ \forall x.\ f(x) = C \cdot g(x)$
- Funkcijam ψ_C pravimo potenciali.

Definicija in izrek prihajata iz [1].

3 Skupna gostota

Kot prej naj bo $G = (V, E)$ in za $uv \in E$ naj bodo $d_{u,v}$ fiksne izmerjene razdalje. Iščemo take vektorje $\underline{x}_V$. ki minimizirajo kvadrate napak med izmerjenimi in predvidenimi razdaljami. Minimiziramo torej t. i. funkcijo energije:

$$E(\underline{x}_V) = \sum_{uv\in E} \left(|x_u - x_v| - d_{u,v}\right)^2$$

V to vsoto vključimo le enega od členov uv in vu. Funkcijo energije lahko komponiramo s strogo padajočo funkcijo e^{-x}, zdaj maksimiziramo

$$P(\underline{x}_V) = exp\left[-\sum_{uv\in E} \left(|x_u - x_v| - d_{u,v}\right)^2\right] = \prod_{uv\in E} e^{-(|x_u - x_v| - d_{u,v})^2}.$$

Če zdaj definiramo potenciale $\psi_{u,v}(x_u, x_v) = e^{-(|x_u - x_v| - d_{u,v})^2}$ dobimo

$$P(\underline{x}_V) = \prod_{uv\in E} \psi_{u,v}(x_u, x_v)$$

Seveda je vsak par povezanih vozlišč klika grafa. Če dodatno za vse klike C, ki niso par vozlišč definiramo $\psi_C(x_C) = 1$ lahko zapišemo:

$$P(\underline{x}_V) = \prod_{C \text{ klika } G} \psi_C(\underline{x}_C)$$

Dobljena funkcija P je zvezna, strogo pozitivna in omejena navzgor z 1. Če vrednosti $\underline{x}_V$ omejimo na množico A s končnim volumnom, npr. $A = B(0, R), R > 0$, bo integral $I = \int_A P(\underline{x}_V)\, d\underline{x}_V$ neničeln in omejen navzgor z volumnom A. Definiramo lahko torej funkcijo $\tilde{P}(\underline{x}_V) = P(\underline{x}_V)/I$, katere integral po A bo enak 1. Funkcija $\tilde{P}$ zadošča vsem pogojem gostote zvezno porazdeljenega slučajnega vektorja, zato lahko definiramo $\underline{X}_V$ kot slučajni vektor z gostoto $\tilde{P}$. Lege naprav, ki minimizirajo kvadrate napak, so zdaj maksimumi gostote $f_{\underline{X}_V}(\underline{x}_V)$.

Označimo V_a kot množico agentov (neznane lege) in V_s kot množico sider (znane lege). Radi bi našli take lege agentov $\underline{x}_{V_a}$, ki minimizirajo kvadrate napak upoštevajoč znane lege sider. Če označimo $\underline{y}_{V_s}$ fiksne znane lege sider, želimo minimizirati $E\left(\underline{x}_{V_a}, \underline{y}_{V_s}\right)$. To je ekvivalentno maksimiziranju pogojne gostote $f_{\underline{X}_{V_a}|\underline{X}_{V_s}}\left(\underline{x}_{V_a}|\underline{X}_{V_s} = \underline{y}_{V_s}\right)$, saj je

$$f_{\underline{X}_{V_a}|\underline{X}_{V_s}}\left(\underline{x}_{V_a}|\underline{X}_{V_s} = \underline{y}_{V_s}\right) = f_{\underline{X}_V}\left(\underline{x}_{V_a}, \underline{y}_{V_s}\right)/f_{\underline{X}_{V_s}}\left(\underline{y}_{V_s}\right) \propto f_{\underline{X}_V}\left(\underline{x}_{V_a}, \underline{y}_{V_s}\right) = \tilde{P}\left(\underline{x}_{V_a}, \underline{y}_{V_s}\right)$$

Ker smo gostoto $\underline{X}_V$ definirali kot produkt potencialov, bo po Hammersley-Cliffordovem izreku G in $\underline{X}_V$ Markovsko polje. Iskanje najbolj verjetne vrednosti v Markovskem polju je problem, ki se ga pogosto rešuje z razširjanjem zaupanja.

V tej izpeljavi so bili potenciali izbrani po vzoru [2] in [3].

4 Razširjanje zaupanja v drevesih

Želimo torej izračunati maksimum

$$f_{\underline{X}_{V_a}|\underline{X}_{V_s}}\left(\underline{x}_{V_a}|\underline{X}_{V_s} = \underline{y}_{V_s}\right) \propto \prod_{\substack{\{u,v\}\in E \\ u,v \text{ agenta}}} \psi_{u,v}(x_u, x_v) \prod_{\substack{\{u,v\}\in E \\ u \text{ agent}, v \text{ sidro}}} \psi_{u,v}(x_u, y_v).$$

To bomo dosegli z razširjanjem zaupanja, cilj te metode so robne gostote

$$f_{X_v|\underline{X}_{V_s}}\left(x_v|\underline{X}_{V_s} = \underline{y}_{V_s}\right).$$

Če bi robno gostoto računali z integriranjem bi morali razrešiti $|V_a| - 1$ dimenzionalni integral, kar je neugodno.

Definiramo sporočilo med sosednjima vozliščema $uv \in E$, ki nista oba sidri, kot v [1], [2], [3]:

$$\text{Za } u, v \in V_a\text{: } \mu_{u\to v}(x_v) := \int_{X_u} \psi_{u,v}(x_u, x_v) \prod_{t\in N(u)\setminus v} \mu_{t\to u}(x_u)\, dx_u,$$

$$\text{za } u \in V_s \text{ in } v \in V_a\text{: } \mu_{u\to v}(x_v) := \psi_{u,v}(y_u, x_v).$$

Poudarili bi, da so sporočila za razliko potencialov nesimentrična. Sporočilo $\mu_{u\to v}$ je funkcija lege prejemnika x_v. Razumemo jo kot ocena pošiljatelja u o verjetnosti lege x_v glede na informacije, ki jih ima u na voljo. Produkt $\prod_{t\in N(u)\setminus v}\mu_{t\to u}(x_u)$ vseh sporočil, ki jih prejme u razen tistega od v oceni kako verjetna je lega x_u. Potencial $\psi_{u,v}(x_u,x_v)$ pa za obravnavano lego x_u oceni lego x_v z uporabo njune izmerjene medsebojne razdalje. Integriranje razumemo kot računanje robne gostote, tako da dobimo funkcijo, ki je odvisna samo od lege x_v.

Definicija sporočila je rekurzivna, za izračun sporočila potrebujemo druga sporočila. Naravno se je vprašati, če se ta rekurzija kdaj konča, torej če lahko sploh izračunamo vsa sporočila.

Trditev 4.1. *Naj bo $G=(V,E)$ drevo, $|V|=n$, $|E|=m$, $V=V_a\cup V_s=\{v_1,\dots,v_n\}$ in sporočila definirana kot zgoraj. Obstaja taka ureditev $(v_{i_1},v_{j_1}),\dots,(v_{i_{2m}},v_{j_{2m}})$ urejenih parov (v_i,v_j), ki ustrezajo povezavam $v_iv_j\in E$, da lahko na vsakem koraku $k=1,\dots,2m$ izračunamo sporočilo $\mu_{v_{i_k}\to v_{j_k}}$ z že izračunanimi sporočili.*

Dokaz. Delamo indukcijo po m.
Najprej $m=1$.
Sporočilo se izračuna bodisi tako, da ustavimo y_u v $\psi_{u,v}$, kar lahko vedno naredimo, bodisi tako, da se integrira $\int_{X_u}\psi_{u,v}(x_u,x_v)\prod_{t\in N(u)\setminus v}\mu_{t\to u}(x_u)\,dx_u$, a je množica v produkti prazna, zato je integral enak $\int_{X_u}\psi_{u,v}(x_u,x_v)\,dx_u$, kar lahko izračunamo.
Predpostavimo, da trditev velja za m in pokažimo, da velja za $m+1$.
Naj bo $s\in V$ list in $v\in V$ njegov edini sosed in naj bo $(v_{i_1},v_{j_1}),\dots,(v_{i_{2m}},v_{j_{2m}})$, ureditev v podgrafu $G\setminus s$, ki obstaja po indukcijski predpostavki. Trdimo da je ureditev $(s,v),(v_{i_1},v_{j_1}),\dots,(v_{i_{2m}},v_{j_{2m}}),(v,s)$ dobra za cel graf G. Najprej izračunamo $\mu_{s\to v}(x_v)=\int_{X_s}\psi_{s,v}(x_s,x_v)\,dx_s$. Za izračun sporočil med $V\setminus s$ ne potrebujemo $\mu_{v\to s}$, saj $s\notin V\setminus s$ (v izračunu sporočila, katerega pošiljatelj je $u\in V$, potrebujemo le taka sporočila, katerih prejemnik je u). Lahko torej po vrstnem redu, ki obstaja zaradi I. P., izračunamo vsa sporočila med vozlišči v $V\setminus s$. Ko smo končali s tem, imamo izračunana vsa ostala sporočila, zato lahko izračunamo tudi $\mu_{v\to s}$.

□

Definiramo še zaupanje $M_u(x_u)\coloneqq\prod_{t\in N(u)}\mu_{t\to u}(x_u)$, ki je produkt vseh prejetih sporočil. Ta po prejšnjem načinu razmišljanja predstavlja oceno lege x_u glede na podatke vseh njegovih sosedov. Pokažimo, da je ta način razmišljanja utemeljen.

Izrek 4.2. *Naj veljajo enake predpostavke kot v trditvi 4.1 in zaupanju definiranem kot zgoraj. Za vsak $v\in V_a$ je zaupanje proporcionalno robni gostoti.*

$$M_v(x_v)\propto f_{X_v|\underline{X}_{V_s}}\left(x_v|\underline{X}_{V_s}=\underline{y}_{V_s}\right).$$

Dokaz. V osnutku bomo dokazali le za $V_s=\emptyset$, sicer velja v splošnem. Fiksiramo $v\in V$.
Vemo, da so poti v drevesu enolične. Za vsak $u\in V\setminus v$ lahko torej definiramo $T^u=(V^u,E^u)$ kot poddrevo vozlišč, za katere pot od v do njih vodi skozi u (vključimo $u\in V^u$) in to obravnavamo kot drevo s korenom u. Za $u\neq v$

definiramo u^+ kot prvi sosed u na poti od u do v (lahko je tudi $u^+ = v$). Velja $u^+ \notin V^u$. Trdimo, da je za $u \neq v$

$$\mu_{u \to u^+}(x_{u^+}) = \int_{\underline{X}_{V^u}} \psi_{u,u^+}(x_u, x_{u^+}) \prod_{ts \in E^u} \psi_{t,s}(x_t, x_s)\, d\underline{x}_{V^u}.$$

Produkt zgoraj razumemo tako, da se v njem pojavi le eden od $\psi_{t,s}$ in $\psi_{s,t}$. To bomo pokazali z indukcijo po globini r drevesa T^u s korenom u.
Najprej $r = 0$.
Tedaj je $V^u = \{u\}$, povezav pa ni. Trditev dobi obliko $\mu_{u \to u^+}(x_{u^+}) = \int_{X_u} \psi_{u,u^+}(x_u, x_{u^+})\, dx_u$, kar je definicija sporočila, če je pošiljatelj list.
Predpostavimo, da velja za vsa poddrevesa globine r in dokažimo, da velja za $r + 1$.
Po definiciji sporočila je

$$\mu_{u \to u^+}(x_{u^+}) = \int_{X_u} \psi_{u,u^+}(x_u, x_{u^+}) \prod_{t \in N(u) \setminus u^+} \mu_{t \to u}(x_u)\, dx_u.$$

Za $t \in N(u) \setminus u^+$ so pa poddrevesa T^t velikosti r, zato lahko uporabimo indukcijsko predpostavko. Opazimo tudi $t^+ = u$ za $t \in N(u) \setminus u^+$ in preuredimo zgornjo enačbo v

$$\int_{X_u} \psi_{u,u^+}(x_u, x_{u^+}) \prod_{t \in N(u) \setminus u^+} \int_{\underline{X}_{V^t}} \psi_{t,u}(x_t, x_u) \prod_{ks \in E^t} \psi_{k,s}(x_k, x_s)\, d\underline{x}_{V^t} dx_u.$$

Velja pa $V^u = \{u\} \cup \bigcup_{t \in N(u) \setminus u^+} V^t$ in $E^u = \bigcup_{t \in N(u) \setminus u^+} (E^t \cup tu)$ in vse te unije so disjunktne. Lahko torej integrale zgoraj združimo v

$$\int_{\underline{X}_{V^u}} \psi_{u,u^+}(x_u, x_{u^+}) \prod_{ts \in E^u} \psi_{t,s}(x_t, x_s)\, d\underline{x}_{V^u}.$$

To je pa formula, ki jo želimo pokazati.

Po dokazanem bo

$$M_v(x_v) = \prod_{u \in N(v)} \mu_{u \to v}(x_v) = \prod_{u \in N(v)} \int_{\underline{X}_{V^u}} \psi_{u,v}(x_u, x_v) \prod_{ts \in E^u} \psi_{t,s}(x_t, x_s)\, d\underline{x}_{V^u}.$$

Podobno kot zgoraj razmislimo, da velja

$$\{v\} \cup \bigcup_{u \in N(v)} V^u = V \quad , \quad \bigcup_{u \in N(v)} (E^u \cup uv) = E.$$

Zato lahko integrale in produkte v zgornji enačbi združimo v

$$\int_{\underline{X}_{V \setminus v}} \prod_{ts \in E} \psi_{t,s}(x_t, x_s)\, d\underline{x}_{V \setminus v}$$

Če se spomnimo definicije slučajne spremenljivke $\underline{X}_V$, je

$$\int_{\underline{x}_{V \setminus v}} \prod_{ts \in E} \psi_{t,s}(x_t, x_s)\, d\underline{x}_{V \setminus v} \propto \int_{\underline{x}_{V \setminus v}} f_{\underline{X}_V}(\underline{x}_V)\, d\underline{x}_{V \setminus v}.$$

Na desni ni pa nič drugega kot

$$f_{X_v}(x_v).$$

$\square$

Literatura

[1] C. Bishop, *Pattern recognition and machine learning*, Springer, New York, 2006. (poglavji 8.3 in 8.4)

[2] V. Savic in S. Zazo, *Belief propagation techniques for cooperative localization in wireless sensor networks* iz *Handbook of position location: theory, practice, and advances*, Wiley-IEEE Press, 2019.

[3] A. Ihler, *Inference in sensor networks: graphical models and particle methods*, doktorsko delo, Department of electrical engineering and computer science, Massachusetts institute of technology, 2005.