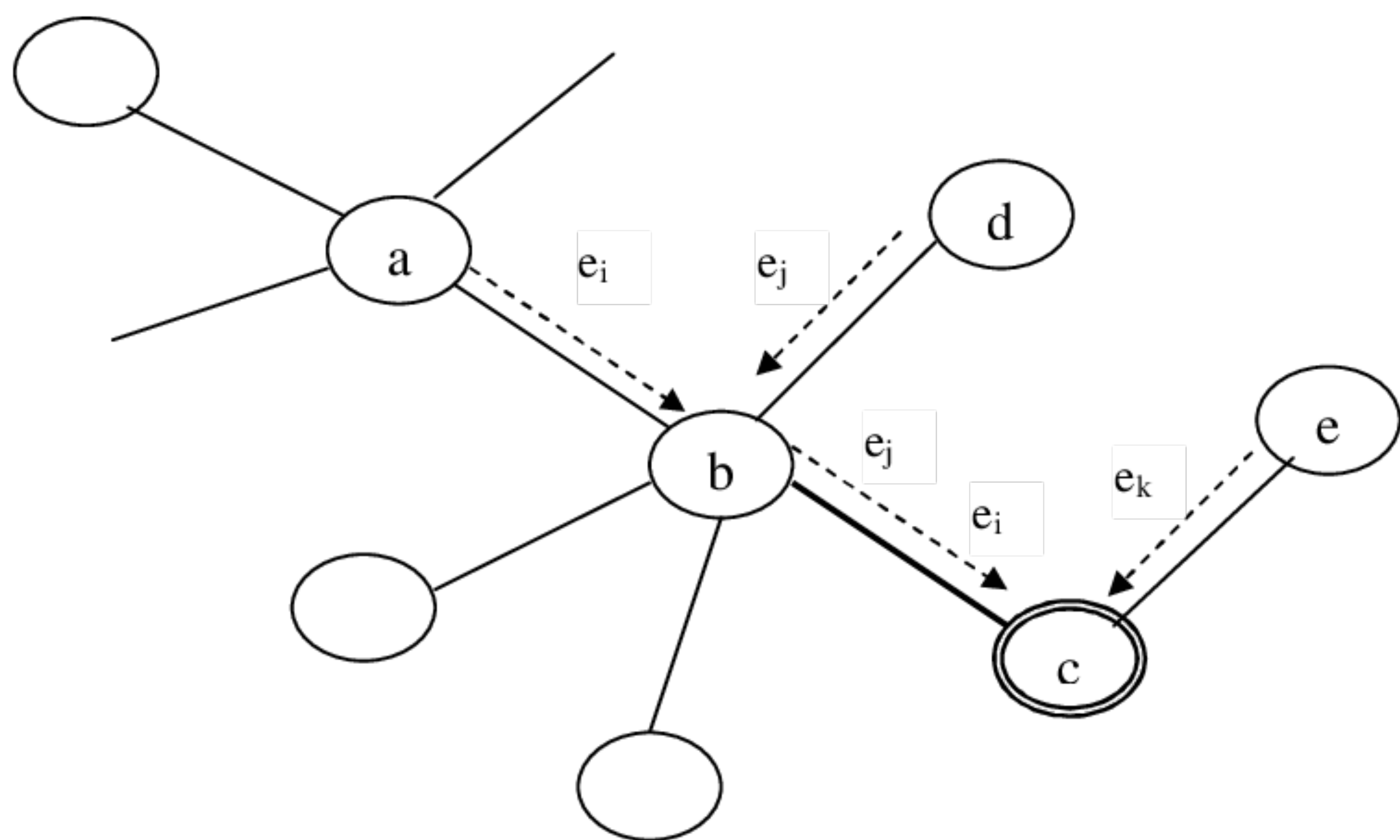


Vahvistusoppimisen sovellukset langattomissa anturiverkoissa ja sen vaikutukset tuotekehityksessä

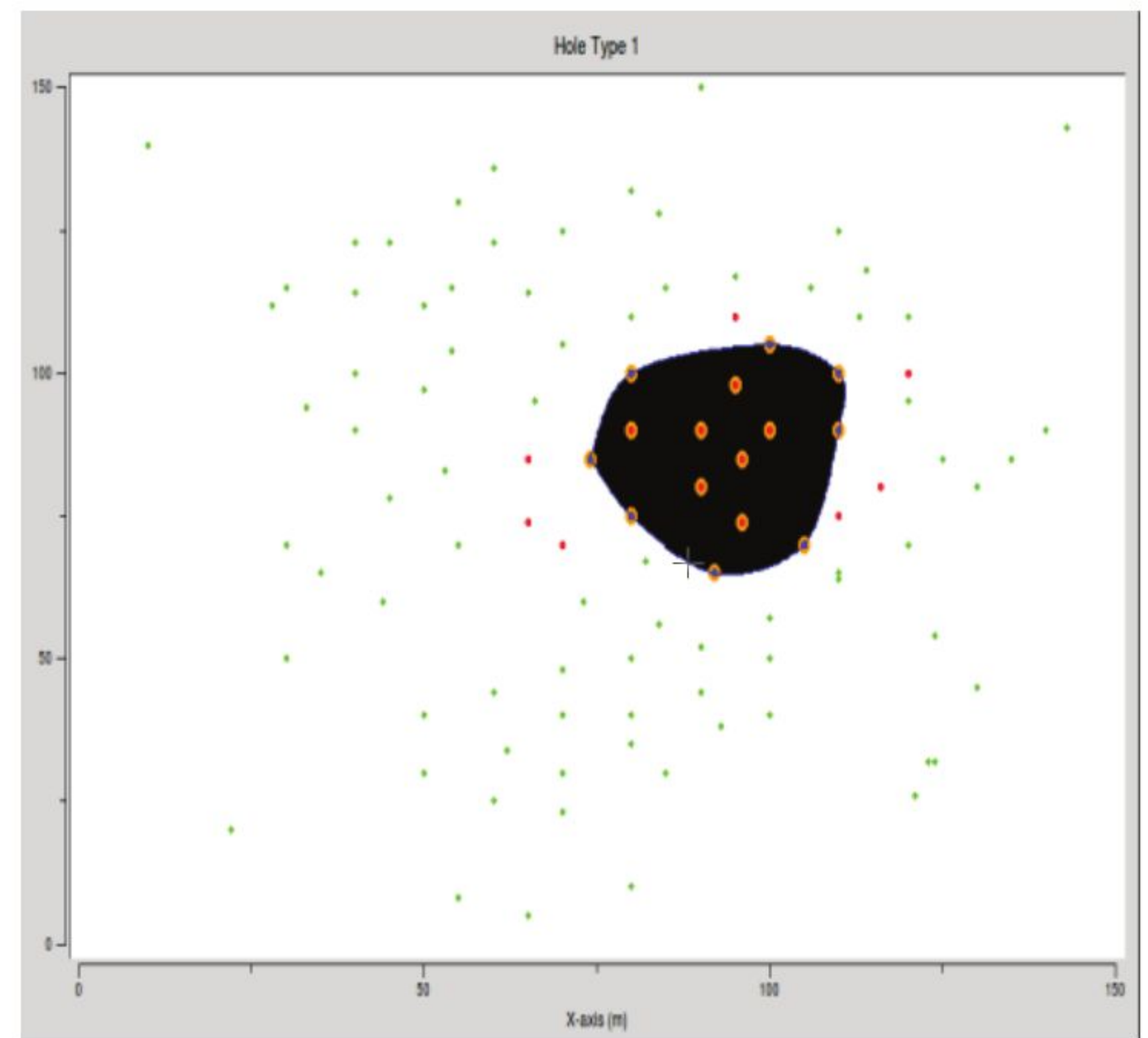
Jonas Nikula

Avainsanat ja keskeiset käsitteet

- **Anturiverkot** koostuvat toisiinsa linkitetyistä antureista. Nykyään puhutaan lähes aina langattomista verkoista (Wireless Sensor Network, WSN). Anturit voivat pysyä paikallaan tai olla liikkuvia, esimerkiksi robotteja tai droneja. Antureiden keräämä data voidaan käsitellä joko keskitetysti, tai sitten hajautetusti. [1]
- **Vahvistusoppiminen** (reinforcement learning) on koneoppimisen alalaji, josta käytetään myös nimitystä “robottioppiminen”, koska se on sen yleisin sovelluskohde. Vahvistusoppimisessa koneoppimismallia muokataan ympäristön ja tilanteen mukaan. Vahvistusoppiminen perustuu oppijan (robotti, tekoäly tms) palkitsemiseen tai rankaisemiseen riippuen siitä kuinka hyvin se suorittaa sille annetun tehtävän (reitit etsiminen, pelin pelaaminen tms). Palkkion ja rankaisun tulisi vaikuttaa oppijan toimintaan niin että sen suoriutuminen tehtävästä paranee. [2]



Esimerkki datan reitityksestä verkossa joka käyttää “path reinforcement algoritmia”. Ympyrät ovat verkon nodeja, e:t toisiinsa liittyvien tapahtumien dataa ja katkoviivat datan reittejä verkossa. Kuten nodesta b nähdään, algoritmi lähettää toisiinsa liittyvien tapahtumien dataa eteenpäin samaa reittiä pitkin. [4]



Simulaatiokuva anturiverkosta ja “reiästä” jonka reititysalgoritmi on havainnut. “Reikä” johtuu sen sisällä olevien anturien “kuolemasta”, joka tarkoittaa että ne eivät syystä tai toisesta pysty lähettämään tai vastaanottamaan dataa, ainakaan tarpeeksi hyvin. [3]

Anturiverkkojen reitityksen haasteet ja menetelmät

- Modernit anturiverkot voivat olla liikkuvia ja ne toimivat usein vaikeissa ympäristöissä
- Modernit anturit keräävät myös entistä enemmän dataa
- Haasteena: Datan reititys tehokkaasti, luotettavasti ja nopeasti sinne missä sitä tarvitaan
 - Vaikka verkon nodet liikkuvat tai niitä kuolisi
- Ratkaisuna vahvistusoppiminen?
 - Tarkastelluissa artikkeleissa menetelmänä pääasiassa Q-learning
 - Ei vaadi mallia ympäristöstä
 - Laskee toiminnan hinnan ja hyödyn odotusarvon
 - Oppii toiminnastaan päivittämällä odotusarvoja (Q-arvo) toteutuneen hinnan ja hyödyn mukaan
- [3] artikkelissa keskitytään löytämään nopein ja luotettavin reitti siirtää dataa nielle.
- Q-arvoon vaikuttaa pääasiassa linkin laatu
- [4] artikkelissa pyritään optimoimaan “datafuusiota”
 - datafuusio tarkoittaa yhteenliittyvän datan yhdistämistä
 - Algoritmin perusidea on lähettää samankaltainen data samalle naapurille
 - Samalla kuitenkin opitaan naapureiden luotettavuus

Lähteet:

- [1] Chee Yee Chong ja Srikanta P. Kumar. “Sensor networks: Evolution, opportunities, and challenges”. Proceedings of the IEEE 91.8 (2003), s. 1247–1256. ISSN: 0018 9219. DOI: 10.1109/JPROC.2003.814918.
- [2] Leslie Pack Kaelbling, Michael L. Littman ja Andrew W. Moore. “Reinforcement learning: A survey”. Journal of Artificial Intelligence Research 4 (1996), s. 237–285. ISSN: 10769757. DOI: 10.1613/jair.301. arXiv: 9605103 [cs].
- [3] Anju Arya, Amita Malik, and Sanjay Kumar. 2015. A Routing Protocol for Detecting Holes in Wireless Sensor Networks with Multiple Sinks. In Proceedings of the Third International Symposium on Women in Computing and Informatics (WCI '15), Indu Nair (Ed.). ACM, New York, NY, USA, 103-108. DOI: <http://dx.doi.org/10.1145/2791405.2791480>
- [4] Bin Yu, Paul Scerri, Katia Sycara, Yang Xu, and Michael Lewis. 2006. Scalable and reliable data delivery in mobile ad hoc sensor networks. In Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems (AAMAS '06). ACM, New York, NY, USA, 1071-1078. DOI: <http://dx.doi.org/10.1145/1160633.116082>