

Atviro kodo tyrimai: Tiekimo grandinės ir saugumas

Studies of Open Source Software: Software Supply Chains and Security

Atvirojo kodo programinė įranga (OSS) yra ypatingos svarbos infrastruktūra, būtina visuomenei ir sukurta savanorių bei prisidedančių įmonių pastangomis. Perėjimą per pastaruosius 20 metų nuo įdomybių iki ypatingos svarbos infrastruktūros lėmė OSS tiekimo grandinių lankstumas, skatinantis atviras inovacijas: bet kurį šaltinio kodą galima patikrinti, iš jo išmokti, nukopijuoti į bet kurį kitą projektą, pranešti apie problemas ir pasiūlyti ar pateikti pataisymus. Spartus augimas, didžiulis mastas ir didžiulis sudėtingumas, atsirandantis dėl decentralizuoto sprendimų priėmimo, apsunkina būsimą plėtrą dėl pervargusių savanorių, matomumo stokos ir rizikos, susijusios su priežiūros stoka, nepataisytu programų pažeidžiamumu. Laimei, OSS atvirumas suteikia galimybę įvertinti įvairių tipų tiekimo grandines, nustatyti atitinkamas rizikas ir sukurti metodus, kaip jas sumažinti. Studijų sritis iš esmės yra daugiadalykė, apimanti tiekimo grandines reprezentuojančių socialinių ir techninių tinklų tyrimus, didelius duomenis, skirtus šimtų milijonų projektų ir dešimčių milijardų failų duomenims saugoti ir užklausoms atlikti, bei mašininio mokymosi ir dirbtinio intelekto metodus, reikalingus kuruoti žemos kokybės duomenys.

Open Source Software provides critical infrastructure through the effort of volunteers and contributing companies that power the internet and multitude of applications the society and enterprises rely on. The transition from a curiosity to critical infrastructure over the last 20 years was driven by the flexibility of OSS supply chains fostering open innovation: any piece of source code could be inspected, learned from, copied to any other project, or depended on, with anyone able to report issues and suggest or provide fixes. The rapid growth, the massive scale, and the extreme complexity resulting from decentralized decision making are straining future development with overworked volunteers, lack of visibility, and risks from lack of maintenance, unfixed bugs and vulnerabilities. Fortunately, the openness in OSS provides opportunities to measure diverse types of supply chains, identify corresponding risks, and develop approaches to mitigate them. The study area is inherently multidisciplinary one involving studies of social and technical networks representing supply chains, big data to store and query data on hundreds of millions of projects and tens of billions of files, and machine learning and artificial intelligence methods needed to curate the often-low-quality data.