

“The method for knowledge acquisition from sensor networks and the Internet of Things using model of augmented perception” – streszczenie pracy

mgr inż. Michał Dyk

Głównym celem pracy było opracowanie metody pozyskiwania wiedzy z sieci sensorowych oraz urządzeń włączonych do Internetu Rzeczy. Urządzenia¹ mogą posiadać różnego rodzaju sensory², przy użyciu których obserwują środowisko³. Podstawowym obszarem zainteresowania są sieci M2M (Machine to Machine) oraz MANET (Mobile Ad-hoc network), w których urządzenia komunikują się bezpośrednio ze sobą, lub w razie braku bezpośredniego połączenia pomiędzy nadawcą a odbiorcą za pomocą skoków (ang. hop by hop). Jednocześnie metoda, bez dodatkowych modyfikacji, może być wykorzystana w rozwiązaniach, w których urządzenia komunikują się ze sobą za pośrednictwem sieci Internet, lub mieszanych (np. klaster sieci M2M połączony dedykowaną bramą z siecią Internet).

Metoda opiera się na koncepcji tzw. „potrzeby informacyjnej”⁴. Oznacza to, że użytkownik (w pracy określany precyzyjnie jako agent kognitywistyczny⁵) określa jaką informacja i w jakiej formie jest potrzebna, natomiast rolą otaczającej go sieci urządzeń jest dostarczenie odpowiedzi na tę potrzebę. W tym celu sieć samoorganizuje się dla wypracowania odpowiedzi na postawione pytanie. Jest to w założeniach inne podejście niż obecnie opracowywane rozwiązania w dziedzinie Data Fusion (jak proces JDL), Big Data i IoT Cloud. Opierają się one na gromadzeniu surowych lub nieco wstępnie przetworzonych danych oraz analizowaniu ich w ramach dedykowanych usług. Niejednokrotnie takie podejście wymaga dużej wiedzy dziedzinowej oraz głębokiej znajomości typów sensorów i przetwarzania sygnałów wyjściowych. Metoda wprowadzona niniejszej pracy przenosi całość analizy do sieci urządzeń. Jest to implementacja koncepcji fog computing (przetwarzanie we mgle) bezpośrednio w sieci.

Motywacja dla takiego postawienia problemu wynika z natury samej informacji. Zawsze to nadawca (w kontekście pracy agent kognitywistyczny) określa co dla niego jest

¹ Formalna definicja urządzenia wprowadzona została w Rozdziale 4.1 pracy (Definicja 4.1.1).

² Formalna definicja sensora została wprowadzona w Rozdziale 4.1 pracy (Definicja 4.1.5).

³ Formalna definicja środowiska obserwowanego przez sensory wprowadzona została w rozdziale 4.2 pracy (Definicja 4.2.20). Opiera się ona na Formalnej Teorii Percepcji przedstawionej w Rozdziale 2.

⁴ Formalna definicja potrzeby informacyjnej przedstawiona jest w Rozdziale 5.2 pracy (Definicja 5.2.1). Opiera się ona na teorii informacji wprowadzonej w Rozdziale 3.

⁵ Formalna definicja agenta kognitywistycznego wprowadzona została w Rozdziale 5.1 (Definicja 5.1.1).

informacją, a co nią nie jest. Wynika to z jego wiedzy dziedzinowej oraz kontekstu. Nawet te same dane mogą dostarczyć różnej informacji bądź nie dostarczyć jej wcale, jeśli odbiorca nie potrafi ich zinterpretować⁶. Z tego powodu celem postawionym przed opracowaną metodą jest dostarczenie odpowiedzi, która jest zrozumiała (jest informacją) dla użytkownika. Ponadto użytkownik zadający pytanie do sieci w postaci potrzeby informacyjnej nie musi mieć wiedzy o jej topologii, rodzajach urządzeń oraz sensorach w które są wyposażone. W myśl opracowanej metody sieć jest autonomicznym systemem ze zdolnością do samoorganizacji na potrzebę udzielenia odpowiedzi na zadane pytanie o informację.

Zrealizowanie celu postawionego w pracy wymagało opracowania szeregu modeli formalnych. Posłużyły one następnie zaprojektowaniu algorytmów zachowania się urządzeń w sieci oraz implementacji środowiska symulacyjnego, które umożliwiło zbadanie w sposób jakościowy i ilościowy opracowanej metody. Pierwszym z modeli opisany został w Rozdziale 2 i jest nim model percepcji. Opisuje on w jaki sposób sensory (w które wyposażone są urządzenia) obserwują zjawiska w swoim otoczeniu. Model ten opiera się na Formalnej Teorii Percepcji⁷, która wprowadza pojęcie i definicję *obserwatora*⁸, jest to podstawowy element procesu percepcji. Dzięki temu modelowi możliwe jest określenie co i w jakiej formie pojedynczy sensor może wiedzieć o otoczeniu. Drugi model opisany został w Rozdziale 3⁹ i jest nim model informacji. Jest on bazą dla późniejszego zdefiniowania potrzeby informacyjnej oraz odpowiedzi na nią. Podstawą modelu jest teoria informacji opracowana przez Keitha Devlina¹⁰. Wprowadza ona pojęcie *infonu*¹¹ jako atomowej jednostki informacji, która stwierdza że pewne obiekty ($a_1 \dots a_n$) znajdują się (bądź nie) w relacji R ¹². Infon posiada formalną definicję, dzięki czemu może być przetwarzany i interpretowany przez urządzenia. Trzecim modelem jest model sieci urządzeń opisany w Rozdziale 4. Model ten oparty jest o teorię grafów i sieci. Definiuje on urządzenia¹³, które posiadają budowę warstwową zgodną ze standardem IEEE P2413¹⁴. Wspomniane warstwy to:

- Warstwa percepcji¹⁵ (wg. standardu IEEE P2413 Properties Layer),
- Warstwa komunikacji¹⁶ (wg. standardu IEEE P2413 Information Exchange Layer),
- Warstwa middleware¹⁷ (wg. standardu IEEE P2413 Function/Method Layer).

⁶ Rozważania na ten temat zawiera Rozdział 3.1.

⁷ Pozycja bibliograficzna nr 15.

⁸ Definicja 2.3.1.

⁹ Rozdział ten zawiera również syntetyczny przegląd różnych teorii informacji, które kształtowały się na przestrzeni rozwoju informatyki.

¹⁰ Pozycja bibliograficzna nr 11.

¹¹ Definicja 3.2.1.

¹² Prosty przykładem jest infon postaci $\langle\langle isOpen, Doors, Room26; now; 1 \rangle\rangle$, który informuje o tym że w obecnej chwili drzwi do pokoju numer 26 są otwarte.

¹³ Definicja 4.1.1.

¹⁴ Draft Standard for an Architectural Framework for the Internet of Things (IoT). Pozycje biblioteczne 32 i 33.

¹⁵ Definicja 4.1.4.

¹⁶ Definicja 4.1.8.

¹⁷ Definicja 4.1.9.

Urządzenia komunikują się ze sobą w sposób bezprzewodowy. Jedno urządzenie może posiadać kilka interfejsów komunikacyjnych¹⁸ (jak ma to miejsce np. w przypadku smartfonów). Jeśli dwa urządzenia znajdują się wzajemnie w zasięgu radiowym traktowane są jako sąsiadujące i mogą się bezpośrednio komunikować. W ten sposób urządzenia tworzą sieć, która modelowana jest jako graf¹⁹. Wraz z ruchem urządzeń topologia sieci może ulegać zmianie. Sam proces komunikacji opisany jest w Sekcji 4.2.3. Pozwala on na zastosowanie różnych algorytmów routingu (np. flooding, SPIN, MEDR). Model sieci uwzględnia zjawiska, które mogą być obserwowane przez urządzenia²⁰ (za pomocą sensorów). W tym zakresie uwzględnia on model percepcji wprowadzony w Rozdziale 2²¹.

W oparciu o modele wprowadzone w Rozdziałach 2, 3 oraz 4 w Rozdziale 5 opisana została metoda pozyskiwania wiedzy z sieci urządzeń. W pierwszym kroku wprowadzone zostało pojęcie *agenta kognitywistycznego*²² (za K. Devlinem). Jest to byt, który posiada, w pewnym zakresie, zdolność do identyfikacji zjawisk oraz obiektów w swoim otoczeniu. Dla przykładu prosty termostat może być traktowany jako agent kognitywistyczny, gdyż „potrafi” zaklasyfikować mierzoną przez siebie temperaturę do stanu ciepło lub zimno. Sam termometr nie może być uznany za agenta kognitywistycznego, gdyż „tylko” mierzy temperaturę i nie poddaje tych obserwacji dalszej analizie. Naturalnie to bardzo proste przykłady. W ogólnym przypadku agent kognitywistyczny może być bardziej złożony. W szczególności może nim być urządzenie, wyposażone w zestaw sensorów, włączone w bezprzewodową sieć sensorową lub Internet Rzeczy. Wymaganiem dla urządzenia, aby uznać je za kognitywistyczne, jest posiadanie odpowiedniej warstwy middleware. W pracy została ona zdefiniowana jako *Cognitive Middleware Layer*²³. Kluczowym jej elementem jest baza wiedzy²⁴, która zawiera podstawowe koncepty²⁵ pozwalające na identyfikację obiektów i zjawisk w obserwowanym środowisku na podstawie odczytów z sensorów.

Zdolność do wyodrębniania obiektów i zjawisk z otoczenia jest podstawą dla opracowanej metody, która opisana została szczegółowo w Rozdziale 5.2. Opiera się ona na, przytoczonej już wcześniej, koncepcji potrzeby informacyjnej²⁶. Potrzeba ta budowania jest w oparciu o model infonów wprowadzony w Rozdziale 3. To użytkownik (bądź urządzenie kognitywistyczne) konstruuje potrzebę, definiując tym samym formę w jakiej chce otrzymać informację zwrotną, czyli odpowiedź na potrzebę informacyjną²⁷. Praca sieci nad przygotowaniem odpowiedzi podzielna jest na etapy:

¹⁸ Definicja 4.1.6.

¹⁹ Definicja 4.2.12.

²⁰ Definicja 4.2.19.

²¹ Związek pomiędzy zjawiskiem w sieci oraz modelem percepcji wyjaśnia diagram przedstawiony na Rysunku 4.7.

²² Definicja 5.1.1.

²³ Definicja 5.1.2.

²⁴ Definicja 5.1.3.

²⁵ Koncepty te znajdują się w ontologii opisanej przez Definicję 5.1.5.

²⁶ Definicja 5.2.1.

²⁷ Definicja 5.2.2.

1. Dystrybucja potrzeby²⁸ – potrzeba zapisana w postaci infonów przesyłana jest do urzędzeń, które mogą udzielić na nią odpowiedzi (posiadają niezbędną wiedzę i sensory) oraz znajdują się na odpowiednim obszarze.
2. Procesowanie potrzeby przez urządzenia²⁹ - każde z urządzeń, które otrzymało wiadomość z potrzebą informacyjną, dokonuje jej analizy i jeśli jest w stanie to zrobić przygotowuje swoją odpowiedź.
3. Uzupelnienie brakującej wiedzy (opcjonalnie)³⁰ – w przypadku, kiedy urządzenie nie ma wiedzy o obiektach których dotyczy potrzeba próbuje ją uzupełnić poprzez pytanie sąsiednich urzędzeń o nową wiedzę.
4. Dystrybucja odpowiedzi na potrzebę³¹ - urządzenia które przygotowały odpowiedź odsyłają ją do nadawcy.

Dzięki zastosowanej notacji infonów przygotowanie potrzeby jest stosunkowo proste. Możliwe jest nawet przekształcenie prostych zdań z języka naturalnego w zestaw infonów (w niektórych przypadkach mogłoby się to dziać automatycznie). Zapis ten jest również formalny, więc może być rozumiany przez komputery, w szczególności inteligentne urządzenia podłączone w bezprzewodową sieć sensorową lub Internet Rzeczy. W opracowanej metodzie ta cecha notacji infonów została wykorzystana. Dzięki temu urządzenia „rozmawiają” ze sobą, korzystając z potrzeby informacyjnej i odpowiedzi na nią. Między innymi wymieniają się wiedzą na temat nowych, nieznanych konceptów w swoich bazach wiedzy. Zastosowanie metody nie jest uzależnione od wykorzystania określonej platformy sprzętowej bądź programowej. Urządzenia muszą jedynie spełniać wymagania postawione agentom kognitywistycznym, to jest posiadać własną bazę wiedzy oraz ontologię zawierającą zbiór podstawowych konceptów umożliwiających identyfikowanie w obserwowanym środowisku obiektów. Opracowana metoda może być traktowana jako rozwiązanie pozwalające zapewnić interoperacyjność urzędzeń w ramach IoT i sieci sensorowych.

Na potrzeby weryfikacji jakościowej i ilościowej opracowanej metody powstał symulator *SenseSim*³², przedstawiony w Rozdziale 6 pracy. Implementuje on modele sieci oraz percepcji wprowadzone w Rozdziałach 4 i 2. Jest to wieloagentowy symulator, który poza aspektami komunikacyjnymi potrafi symulować zachowanie urzędzeń oraz uruchamiane przez nie programy. Ponadto pozwala symulować elementy otoczenia poprzez zmienne w czasie zjawiska, które mogą być obserwowane przez sieć urzędzeń. Dzięki modułowej konstrukcji możliwa jest wymiana każdej z warstw składowych urzędzeń. W szczególności symulowane urządzenia mogą zostać wyposażone w warstwę middleware z zaimplementowanymi

²⁸ Proces dystrybucji potrzeby opisany jest w Definicji 5.2.5.

²⁹ Definicja 5.2.6. opisuje zachowanie urządzenia, które otrzymało wiadomość z potrzebą informacyjną. Definicja 5.2.9 opisuje algorytm przygotowania odpowiedzi na potrzebę przez pojedyncze urządzenie.

³⁰ Definicja 5.2.7.

³¹ Proces dystrybucji odpowiedzi opisany jest w Definicji 5.2.10.

³² Symulator był prezentowany na Światowym Forum Internetu Rzeczy w Mediolanie w 2015 roku. Był również wykorzystywany w pracach Grupy Roboczej NATO IST-147 „Militarne Zastosowania Internetu Rzeczy”, której autor pracy był członkiem w trakcie jej funkcjonowania.

algorytmami opracowanymi w Rozdziale 5. W ten sposób stają się agentami kognitywistycznymi zdolnymi do przetwarzania potrzeby informacyjnej.

Przeprowadzone w ramach pracy eksperymenty zostały opisane w Rozdziale 7. Pokazały, że metoda może być wykorzystana do budowania świadomości sytuacji w czasie rzeczywistym. Zaprezentowany scenariusz dotyczył sytuacji triage-u w trakcie misji wojskowej³³. Dla dowódców kluczowa jest wiedza o stanie fizycznym podległych żołnierzy, gdyż może to mieć wpływ na realizację zadań. Z drugiej strony dowódcy są przytłoczeni zadaniami i zazwyczaj nie są kompetentni do analizowania sygnałów życiowych podległych żołnierzy. Tego rodzaju monitoring ogranicza się nie tylko do identyfikacji sytuacji zagrożenia zdrowia lub życia, ale powinien obejmować także przemęczenie, odwodnienie, nadmierny stres. Aby wykorzystać opracowaną metodę dowódca przygotowuje prostą potrzebę informacyjną. Pozostałą pracę związaną z wnioskowaniem o stanie zdrowia żołnierzy wykonuje sieć urządzeń. Dowódca nie musiał nawet wiedzieć, jakie czujniki noszą żołnierze. Sama metoda dopuszcza sytuację, w której informacje o jednym obiekcie zbiera wiele urządzeń, przy czym każde może uwzględniać inny jego aspekt. W rozegranym scenariuszu każda odpowiedź sieci została przygotowana w czasie rzeczywistym, a użytkownik (w tym przypadku dowódca plutonu) nie musiał wkładać wiele wysiłku w analizę wyników. Wynika to z faktu, iż przygotowując potrzebę informacyjną sam określił co jest dla niego istotną informacją i w jakiej formie ma zostać dostarczona.

Miary ilościowe zebrane podczas eksperymentów dają wgląd w skuteczność metody i możliwości jej zastosowania, zwłaszcza w przypadku mobilnych sieci ad-hoc (MANET). W takich przypadkach moc obliczeniowa urządzeń (np. współczesnych smartfonów lub mikrokomputerów IoT) często stanowi mniejszy problem niż ich możliwości komunikacyjne (np. BLE ogranicza jeden komunikat do 20 bajtów, z kolei LoRa potrafi dać zasięg przekraczający 5km ale transmisja oscyluje wtedy w okolicy 5000 bit/s). Dodatkowo komunikacja bezprzewodowa zużywa dużo energii, a wysoka przepustowość ma ograniczony zasięg. Dlatego metoda była oceniana głównie przez pryzmat dwóch kluczowych czynników: czasu potrzebnego na dostarczenie odpowiedzi oraz ilości danych które sieć musi przetworzyć podczas procesu odpowiadania na potrzeby. Sam czas obsługi potrzeby w urządzeniu może się różnić w zależności od platformy sprzętowo-programowej, dlatego ten czynnik nie był uwzględniany w końcowych wynikach. Poprawność działania algorytmów udzielających odpowiedzi w urządzeniu została potwierdzona w trakcie rozgrywania różnych wariantów scenariusza triage. Wyniki eksperymentów ilościowych pokazały, że nawet dla niskiej przepustowości (5 000 bitów / s – czyli np. dla LoRa) odpowiedź może być wysłana do odbiorcy, w ciągu kilku sekund (w najbardziej złożonych spośród rozpatrywanych scenariuszy). Rozmiar sieci nie ma istotnego wpływu na ilość danych wymienianych przez pojedyncze urządzenia, dzięki czemu metoda skaluje się na sieci o różnym rozmiarze (w sensie liczby połączonych urządzeń) oraz topologii.

³³ Scenariusz ten był prezentowany w trakcie demonstracji przygotowanej przez Grupę Roboczą NATO IST-147 w Oulu podczas „*International Conference On Military Communications And Information Systems*” w 2017 roku.

Metoda pozyskiwania wiedzy jest głównym rezultatem pracy. Jednak w trakcie badań powstały inne wyniki, które można traktować jako pojedyncze produkty. Są to:

- Model Bezprzewodowej Sieci Sensorowej (WSN) - jest kompletnym modelem matematycznym WSN, który obejmuje nie tylko aspekty komunikacyjne, ale także konstrukcję urządzeń, ich zdolności percepcyjne, jak również środowisko zewnętrzne względem sieci, w którym mogą występować zmieniające się w czasie zjawiska (obserwowane przez urządzenia).
- Symulator SenseSim - symulator to implementacja formalnego modelu sieci WSN. Koncentruje się nie tylko na aspektach komunikacyjnych, ale także może symulować zachowanie urządzeń, jak również zjawiska w zmieniającym się środowisku. Ma także możliwość łączenia rzeczywistych urządzeń, dzięki czemu można przy jego użyciu przeprowadzać eksperymenty typu Hardware In The Loop (HWITL).