

prof. dr hab. inż. **Piotr Kulczycki**  
Polska Akademia Nauk, Instytut Badań Systemowych  
Akademia Górniczo-Hutnicza, Wydział Fizyki i Informatyki Stosowanej  
ul. Jaracza 30, 31-216 Kraków  
tel.: 601 314159  
e-mail: kulczycki@ibspan.waw.pl, kulczycki@agh.edu.pl

## RECENZJA

rozprawy doktorskiej

mgr inż. Michał Madera

*„Predykcja defektów w kodzie źródłowym z zastosowaniem rozmytego systemu wnioskującego”*

Recenzowaną rozprawę wykonano pod kierunkiem promotora Pana prof. dr hab. inż. Jacka Kluski oraz promotora pomocniczego Pana dr hab. inż. Macieja Kusego. Przewód doktorski został wszczęty na Politechnice Rzeszowskiej, w dziedzinie nauk inżynieryjno-technicznych, dyscyplinie *Informatyka techniczna i telekomunikacja*, według tzw. starego trybu postępowania.

### 1. Charakterystyka rozprawy

Praca sprawia znakomite wrażenie od pierwszego kontaktu. Szata graficzna godna książki (wydawnictwa uczelnianego), strona edytorska i układ są bezbłędne. Treść jest staranna, z wyraźną dbałością o jakość języka, a to poprzez użycie wyrazów nietrywialnych, czy też unikania powtórzeń wyrazu w sąsiednich zdaniach (np. str. 37<sup>1-2</sup> zamiast „adresujących” można było użyć powszechnego „uwzględniających”, ale wyraz ten został już użyty bezpośrednio po tym). Błędów literowych praktycznie nie ma (dla wykazania swojej czujności podam, że na str. 64<sub>8</sub> zamiast „algorytm” powinno być „algorytmu”, a na 41<sup>12</sup> jest „wy” zamiast „w”). Powyższe powoduje, że czytelnik może z całą przyjemnością przenieść swoją uwagę na merytoryczną stronę przedłożonego tekstu.

Przedmiotem pracy jest synteza procedury, która wskazywałaby w kodzie źródłowym programu fragmenty potencjalnie wadliwe, które mogą skutkować późniejszym błędnym działaniem oprogramowania. Następuje to poprzez dwustanową klasyfikację prowadzoną na rozważanym kodzie (zły *versus* dobry) i – co należy wyraźnie podkreślić – dodatkowo przy wymaganiu możliwie dużej interpretowalności powyższego wyniku, użytecznej dla potrzeb programisty. W tym celu została zaproponowana metoda oparta na regułach rozmytych typu „jeżeli (...) to (...)”, przy czym dane wejściowe mogą być zarówno ciągłe jak i kategoryczne. (Taka hybrydyzacja bywa w praktyce dodatkowym wyzwaniem, wynikłym chociażby z odmiennych aparatów matematycznych – w przypadku atrybutów ciągłych opartych na funkcji gęstości, a w przypadku kategorycznych mierze skoncentrowanej na skończonej liczbie punktów, czyli jakby zaprzeczeniu samego charakteru funkcji gęstości oraz warunków jej istnienia). Do wydobywania metareguł użyto metod genetycznych, ściślej algorytmu programowania ekspresji genów. Dużą uwagę poświęcono ocenie jakości otrzymanego rozwiązania. Zaprojektowany ostatecznie w pracy algorytm GPR został poddany obszernej weryfikacji i porównania z dostępnymi w literaturze alternatywnymi metodami. Ostatecznie okazał się on lepszy, a przynajmniej nie gorszy niż metody alternatywne. Wyróżnia się za to dogodną interpretowalnością wyników. Może także pomóc wskazać aspekty wspomagające kontrolę jakości wytworzonego oprogramowania. Są to własności wyjątkowo przydatne w praktyce zaawansowanego programowania. Statystyczną istotność efektywności algorytmu GPR pozytywnie zweryfikowano na drodze testowania statystycznego.

Recenzowana praca liczy 138 stron tekstu, podzielonego na 7 rozdziałów o klasycznym dla tego typu prac charakterze. Rozdział 1 stanowi wstęp, w drugim przedstawiono wyniki badań bibliograficznych, po czym w trzecim syntetycznie ujęto tezę, cel i zakres rozprawy. Krótki rozdział 4 dostarcza preliminariów metodycznych. Rozdział 5 tworzy kwintesencję pracy – przedstawiono tu syntezę klasyfikatora w postaci systemu decyzyjnego opartego na regułach rozmytych i programowaniu ekspresji genów. Obszerny rozdział 6 przedkłada wyniki badań empirycznych, dyskusję uzyskanych efektów oraz analizę porównawczą. Rozdział 7 stanowi podsumowanie, po którym następuje obszerna bibliografia złożona z 132 trafnie dobranych pozycji, a następnie dwa prawidłowo dobrane krótkie dodatki oraz anglojęzyczne streszczenie.

Szczegółowy wkład naukowo-badawczy Doktoranta został podany na stronie 113 rozprawy. Trudno się z nim nie zgodzić. Jednak w moim odczuciu głównym osiągnięciem jest całościowa synteza wielowątkowego algorytmu, łączącego szereg wartościowych koncepcji, który stanowi realną konkurencję dla istniejących metod. Łączenie i dostosowanie często niepasujących do siebie *a priori* podprocedur potencjalnie sprawia trudne do przewidzenia kłopoty, niewidoczne po pomyślnym zakończeniu dzieła, i z tego powodu nierzadko są one niedoceniane.

## 2. Uwagi krytyczne i dyskusyjne

a.

Przede wszystkim chciałbym poddać pod dyskusję zasadność pierwszego wyrazu tytułu recenzowanej rozprawy: „predykcja”. Jest to synonim słowa „prognozowanie”, czyli wyznaczanie, na podstawie bieżących danych, przyszłej sytuacji w określonym horyzoncie czasowym. Przykładowo, na podstawie dzisiejszego rozkładu ciśnienia, temperatury i wiatru można prognozować jutrzejsze opady. Tymczasem praca dotyczy raczej wskazania tych elementów kodu źródłowego, które grożą nieprawidłowościami w przyszłości, przy braku ścisłego określenia tych nieprawidłowości i – przede wszystkim – bez horyzontu czasowego. Wydaje mi się, że jest to bardziej „wykrycie”, „detekcja”, ale bez czynnika czasu, typowego dla prognozy *vel* predykcji. Nawiązując do powyższego przykładu meteorologicznego, w wyniku zastosowania opracowanej procedury otrzymuje się raczej wskazanie które miejsca produkowanej kurtki należy wzmocnić, aby użytkownik możliwie najmniej ucierpiał na skutek zjawisk atmosferycznych w okresie eksploatacji. To nie jest predykcja.

Zdaję sobie sprawę, że współcześnie – dla przyciągnięcia potencjalnego czytelnika – „klasyfikacja” czasem nazywana jest „predykcją” (jako przewidywaniem etykiety), wzory prowadzące do tego „modelem”, a jeśli dzieje się to na podstawie sukcesywnie napływających danych, to można dodać modne „systemy uczące się” i już z klasycznej „klasyfikacji” mamy „model predykcyjny z użyciem systemów uczących się”. Ładnie brzmi, ale w moim odczuciu niekoniecznie w pracy doktorskiej.

Liczę na miłą i konstruktywną dyskusję w trakcie obrony.

b.

W sekcji 6.4.4 (str. 98-100) Autor przedstawia wyniki testu Wilcoxona, weryfikując hipotezę o przypadkowości przewagi skonstruowanej w recenzowanej pracy procedury względem innych porównywanych algorytmów. Otóż, test Wilcoxona jest tzw. testem istotności, gdzie weryfikowaną hipotezę zerową odrzuca się z założonym prawdopodobieństwem błędu pierwszego rodzaju (odrzuć hipotezy prawdziwej) lub nie odrzuca, bynajmniej nie przyjmując hipotezy alternatywnej (nawet można jej nie formułować), gdyż zachodziłoby niebezpieczeństwo błędu drugiego rodzaju, który w testach istotności jest niekontrolowany. Tak więc ilość informacji uzyskiwanej z testu istotności nie jest duża.

Wartość powyższego prawdopodobieństwa nazywana jest poziomem istotności i w praktyce zwykle przyjmuje się 0,01, (może także 0,02), 0,05, 0,1, ze wskazaniem na 0,05. Tę właśnie wartość przyjęto w recenzowanej dysertacji i dla niektórych algorytmów uzyskano wyniki pozytywne (odrzuć), dla innych neutralne (brak odrzucenia). W sytuacji tak płytkich wniosków, dodatkowe wskazania można było uzyskać przyjmując „ostrzejsze” poziomy istotności

0,01 i 0,02 (co wzmocniłoby wynik pozytywny) lub „słabszy” 0,1, co być może pozwoliłoby uzyskać wynik pozytywny (zamiast neutralnego), aczkolwiek słabiej uwarunkowany.

Liczę na przedstawienie wyników takich badań w trakcie obrony.

c.

Z dużą dozą pozytywnego uśmiechu skonstatowałem konstrukcję rozdziału 2 – badania bibliograficzne zostały tu podane w postaci trzyetapowej sformalizowanej procedury. Zastanawiam się czy czynność tak nieokreślona – z samej swej natury – jak studia bibliograficzne powinny być ujmowane w ramach ścisłej procedury? Czy takie postępowanie nie zawęży możliwości poznania nowego, nieprzewidywalnego (a więc kwintesencji nauki i badań wyprzedzających), spoza założonego *a priori* zakresu?

Liczę na miłą wymianę poglądów w trakcie obrony naukowców-badaczy z dwóch różnych pokoleń, a może nawet epok: przed- oraz internetowej, tej drugiej coraz bardziej poddanej wszechobejmującym procedurom.

d.

Niejednokrotnie Autor myli metrykę z miarą. Otóż, metryka charakteryzuje relację (np. odległość, podobieństwo) między dwoma elementami, miara natomiast dotyczy jednego obiektu (np. pole powierzchni figury, wskaźnik jakości rozwiązania). W tym ostatnim znaczeniu Doktorant używa wzmiankowanego pojęcia, błędnie przywołując termin metryki. Dotyczy to przede wszystkim podrozdziału 4.4 i wszelkich jego konsekwencji, ale dobitnie uwidocznia się przy okazji wzoru (6.18) na stronie 78<sup>2</sup>. Wielkość *MCC* nazywana jest w pierwszej linii miarą, a w siódmej metryką. To nie są pojęcia zamienne. (Zdaję sobie sprawę, że w języku potocznym metryka czasem dotyczy jednego obiektu, np. metryka chrztu, ale w pracy zmatematyzowanej nie powinno to mieć miejsca.)

Ogólnie: przytoczone powyżej uwagi mają charakter drugorzędny oraz dyskusyjny i nie negują generalnie bardzo pozytywnej mojej oceny przedstawionej dysertacji.

### 3. Konkluzja

**Podsumowując uważam, że recenzowana praca doktorska Pana Michała Madera spełnia wymagania stawiane w odpowiednich przepisach rozprawom doktorskim i wobec tego wnioskuję o dopuszczenie Go do dalszych, przewidzianych ustawą, etapów przewodu doktorskiego.**

Biorąc pod uwagę aktualność badanego zagadnienia, sposób i jakość opracowanego rozwiązania oraz staranność redakcyjną **wnoszę o wyróżnienie rozprawy**. Pozostaję z nadzieją, że po obronie będę mógł podtrzymać tę rekomendację.



(prof. dr hab. inż. Piotr Kulczycki)

Kraków/Warszawa, 2 listopada 2022.