

**WYDZIAŁ TELEKOMUNIKACJI, INFORMATYKI
I ELEKTROTECHNIKI**

Bezpieczeństwo danych - projekt

Algorytm Boyera-Moore'a



**POLITECHNIKA
BYDGOSKA**

im. Jana i Jędrzeja Śniadeckich

Wykonali:

Olech Jakub
Pazderski Paweł
Rykała Szymon
Dawid Dojczman

1. Opis Algorytmu

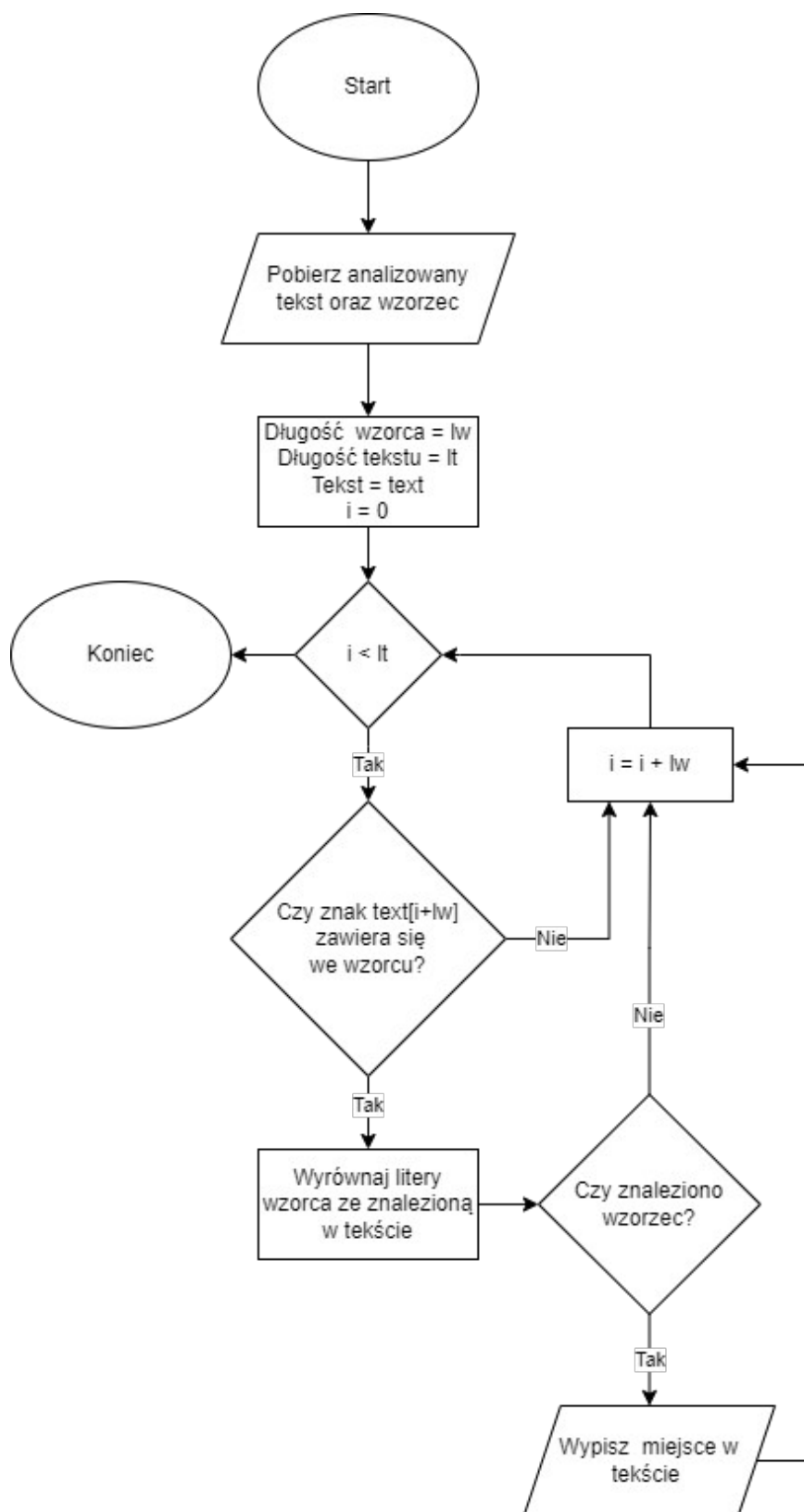
Algorytm Boyera-Moora (B-M) to efektywna metoda wyszukiwania wzorca w tekście, która została opracowana przez Roberta Boyera i J. Strothera Moora w 1977 roku. Algorytm ten jest oparty na heurystyce, która pozwala na szybsze wyszukiwanie wzorca w tekście poprzez wykorzystanie informacji o tym, jak wzorzec pasuje do tekstu. Podstawowa wersja algorytmu Boyera-Moora rozpoczyna porównywanie tekstu i wzorca od ostatniego znaku wzorca.

Pełny algorytm B-M korzysta z dwóch heurystyk:

- Heurystyka good suffix shift jest wykorzystywana w przypadku, gdy pierwsze X znaków wzorca pokrywa się z tekstem, ale reszta wzorca już nie. W takim przypadku algorytm sprawdza, czy w pozostałej części wzorca występuje taki sam ciąg znaków jak na końcu pasującej części wzorca. Jeśli tak, to algorytm przesuwca wzorzec o tę samą liczbę pozycji, aby pasująca część wzorca pokrywała się z tekstem. Dzięki temu unikamy zbędnych porównań między wzorcem a tekstem i przyspieszamy proces wyszukiwania.
- Heurystyka bad-character shift - polega na tym, że jeśli występuje niezgodność między wzorcem a tekstem na pozycji i-tej, to algorytm przesuwca wzorzec w prawo o taką liczbę pozycji, aby ostatnie wystąpienie znaku na lewo od pozycji i w wzorcu pokrywało się z odpowiadającym mu znakiem w tekście. Dzięki temu unikamy zbędnych porównań między wzorcem a tekstem i przyspieszamy proces wyszukiwania.

Algorytm Boyera-Moora i jego warianty są nadal szeroko stosowane w dzisiejszych aplikacjach i programach. Pomimo pojawienia się nowych metod i algorytmów do wyszukiwania wzorców w tekście, algorytm Boyera-Moora pozostaje jednym z najskuteczniejszych i najwydajniejszych sposobów na szybkie wyszukiwanie wzorca w tekście.

2. Schemat blokowy



Rys.1. Schemat blokowy algorytmu Boyera-Moora