

Cyfrowe znaki wodne wspomogą walkę z fake newsami

3 listopada 2021

Na wojnę z fake newsami idą tym razem... specjaliści od ukrywania treści. Mają pomysł, jak dzięki m.in. cyfrowym znakom wodnym i sztucznej inteligencji wskazywać, które materiały foto lub wideo zostały edytowane.



Polityk mówi na filmie rzeczy, których nikt mądry by nie powiedział. Aktor pojawia się w filmie, w którym wydawałoby się, że nie miał możliwości zagrać. Stylizacja sławnej modelki łądząco przypomina zdjęcia mało znanej instagramerki... łatwość w dostępie do technologii deep fake czy narzędzi do edycji obrazów sprawia, że coraz trudniej odróżnić materiał oryginalny od tego, który został cyfrowo przerobiony i wprowadza odbiorców w błąd.

Aby zapobiegać takiej manipulacji opinii publicznej, naukowcy z Japonii, Hiszpani i Polski zjednoczyli siły w ramach projektu DISSIMILAR (Detection of fake newS on SocIal MedIa pLATfoRms), aby naprawdę wspomóc prawdę. O projekcie poinformowali przedstawiciele Wydziału Elektroniki i Technik Informacyjnych Politechniki Warszawskiej (WEiT PW) w

przesłanym PAP komunikacie.

„Rezultatem naszego projektu DISSIMILAR będzie przeznaczona dla mediów społecznościowych platforma pozwalająca na automatyczną weryfikację prawdziwości informacji. Jej działanie będzie opierać się na algorytmach cyfrowego znakowania wodnego treści multimedialnych oraz mechanizmów opartych na sztucznej inteligencji i uczeniu maszynowym. Zweryfikują one automatycznie, czy dany materiał był wcześniej poddany celowej modyfikacji czy nie” – podkreśla prof. Wojciech Mazurczyk z Instytutu Informatyki WEiTI PW kierujący polskim zespołem w projekcie DISSIMILAR. I dodaje: „Nasze prace związane z ukrywaniem informacji pozwalają więc nie tylko na zabezpieczenie oryginalnych komunikatów, ale również znacząco wpływają na wykrycie pojawiających się fałszywych treści”.

Cyfrowe znakowanie wodne to zestaw innowacyjnych technik pozwalający na szybkie sprawdzenie, czy dany materiał multimedialny (tekst, zdjęcie, nagranie czy film) jest oryginalny czy zmodyfikowany. Co więcej dzięki metodzie tej można też wskazać pierwotne źródło pochodzenia otrzymanej informacji.

Rozwiązanie polega na osadzeniu niewidocznych i niesłyszalnych znaków wodnych w chronionych treściach. Usunięcie tych znaków będzie jednak natychmiast dostrzegalne. Dla twórców treści ma to być zabezpieczenie i ochrona ich dzieł przed niechcianymi modyfikacjami.

W ten sposób cyfrowe znakowanie wodne ma stać się narzędziem w walce z dezinformacją i daje możliwość rozwiązania problemów ochrony praw autorskich w sieci. Ma też być odpowiedzią na wyzwania związane z uwierzytelnianiem treści.

Jak podkreślono w komunikacie WEiTI PW, cyfrowe znakowanie wodne nie zawsze pozwala na wykrycie modyfikacji materiału

multimedialnego. Pełną skuteczność dają dopiero odpowiednie narzędzia informatyczne do wykrywania wszelkiego rodzaju modyfikacji, które efektywnie działają nie tylko w najpopularniejszych platformach mediów społecznościowych, ale także zapewniają wysoki poziom pracy w coraz popularniejszych komunikatorach internetowych.

Ze względu na udział trzech różnorodnych partnerów z całego świata: Japonii (Okayama University), Hiszpani (Fundació per a la Universitat Oberta de Catalunya) i Polski (WEiTI PW) projekt DISSIMILAR swoim zakresem obejmuje również badania społeczne, w tym analizę czynników kulturowych związanych z fake newsami. Wyniki będą obejmować całe spektrum zagadnień tworzenia, wykrywania oraz radzenia sobie z dezinformacją w szeroko pojętej opinii publicznej.

Zdjęcie: [pixel2013](#) (CC0)

Źródło: NaukawPolsce.PAP.pl