Editorial

Lorna Uden

Faculty of Computing Engineering and Sciences (FCES),
School of Computing,
Staffordshire University,
College Road, Stoke-on-Trent, Staffordshire ST4-2DE, UK
Email: L.uden@staffs.ac.uk

Welcome to Vol. 14 No. 2 issue of *IJWET*. There are four papers in this issue.

The first paper is 'Semantic lifting and reasoning on the personalised activity big data repository for healthcare research', by Hong Qing Yu and Feng Dong. The authors of this paper developed an innovative framework that manages and integrates multiple health-related data resources from wearable sensors, mobile and web applications. The aim of the research is to form an efficient backend platform and technology packages for mining personalised health knowledge which will exert influence on the future direction of people's self-care empowerment, disease prevention and importantly promote better lifestyles.

The framework applies the hybrid database architecture of NoSQL and RDF repositories with introductions for semantic oriented data mining and knowledge lifting algorithms. The activity stream data is collected through Kafka's big data processing component. The motivation of the research is to enhance the knowledge management, discovery capabilities and efficiency to support further accurate health risk analysis and lifestyle summarisation. However, it is difficult to determine whether the project is successful because there is a lack of evaluation, conclusion and future work included.

The second paper is 'DWSpyder: a new schema extraction method for a deep web integration system', by Yasser Saissi, Ahmed Zellou and Ali Adri. According to these authors, the deep web is a huge part of the web that is not indexed by search engines. The deep web sources are accessible only through their associated access forms. In this paper, the authors described their method, DWSpyder, for extracting the deep web source schema description required to implement a web integration system. To implement web integration system, it is necessary to know the schema description of each web source. These authors have developed the DWSpyder method to extract the schema. The extracted schema is the key information used to build the catalogue needed by a web integration system.

The DWSpyder method starts with a static analysis of the deep web source access forms to extract the first elements of the associated schema description. The second step is a dynamic analysis of these access forms using queries to enrich the schema description. The DWSpyder method also uses a clustering algorithm to identify the possible values of deep web form fields with undefined sets of values. All the information extracted is used by DWSpyder to generate automatically deep web source schema descriptions. The main problem of this research is the lack of evaluation of the method to validate its uses.

The third paper is 'Impact of replica placement-based clustering on fault tolerance in grid computing', by Rahma Souli-Jbali, Minyar Sassi Hidri and Rahma Ben Ayed. This paper presents an approach to optimise fault tolerance techniques for grid systems. The proposed approach based on clustering the nodes into clusters based on dynamic data replication, which reduces the response time and improves the use of storage space and network resources. According to these authors, the aim of the research is to propose a fault tolerance protocol, which within a cluster the system uses message logging, and between clusters use coordinated checkpointing.

These authors argue that results show the efficiency of the proposed protocol compared with similar protocols in terms of recovery time, the number of rollback processes, the number of markers and the impact of using message-logging. The results also show many significant reductions in recovery time by more than 70% and of the number of processes in rollback which reach 90%. More empirical studies are needed to validate the method.

The final paper is 'Anomaly detection in the web logs using user-behaviour networks', by Jingwen You, Xiaojuan Wang, Lei Jin and Yong Zhang. In this paper, the authors proposed user behaviour network model for anomaly detection in web logs. Based on user, by integrating the network structure and the characteristic of anomalous users, they propose five indicators to identify the anomalous users and the anomalous logs' behaviour.

According to these authors, the results show that the method gets a better performance on four real web application log datasets, with a total of about four million log messages and one million anomalous instances. They also argue that the results show that not all the anomalous users initiate attacks. What is more, the method proposed can identify not only the users who make an attack but also users who scan or probe the website to prepare for the attacks. More evaluations are needed to validate the claims.