Growing Software: Objective, Methodology, and Technology

GREETINGS and welcome to the third issue of IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS (TCSS) for 2023. The authors are delighted to share some exciting news with our esteemed readership.

First and foremost, the authors are pleased to announce that TCSS has achieved a remarkable milestone in scholarly impact. As per the latest update of the CiteScore Tracker from Elsevier Scopus, released on 5 April 2023, our journal has achieved a historic high CiteScore of 9.9. This is a testament to the outstanding quality of research that the authors have published and the enduring impact of our scholarly contributions in the field of computational social systems.

In addition to this, the authors are also delighted to inform our readers that our proposal for a 1200-page budget increase has been approved by IEEE. This development is significant as it will enable us to accommodate more high-quality manuscripts, ensuring that they are published in a timely and efficient manner. The authors believe that this will not only enhance the visibility and impact of TCSS but also create a more robust platform for scholarly communication and collaboration in our field.

The authors take this opportunity to express our gratitude to our authors, reviewers, editorial board members, and readers for their unwavering support and contribution to the success of TCSS.

After the usual introduction of our 49 regular articles, the authors would like to discuss the topic of growing software: objective, methodology, and technology.

I. SCANNING THE ISSUE

1. "Solving the Team Allocation Problem in Crowdsourcing via Group Multirole Assignment" by *Lu Liang, Jingdong Fu, Haibin Zhu, and Dongning Liu*

The main contributions of this article include the following: 1) the first formalization of the team allocation problem (TAP) in crowdsourcing through extending group multirole assignment (GMRA) and the creation of a composition matrix to express high-order cardinality (HC) constraints of agents and conflict avoiding constraints through the crowdsourcing scenario via GMRA with constraints (GRA+); 2) theoretical proofs of the theorems of the formalized problem, such as a necessary and sufficient condition. They introduce HC and conflicting agents on roles (CARs) constraints and formalize the TAP via GMRA; and 3) a practice solution to the proposed problem based on the IBM ILOG CPLEX optimization package. All the proposed approaches are verified by simulation experiments.

2. "Two-Stage Auction Mechanism for Long-Term Participation in Crowdsourcing" by *Timothy Shin Heng Mak and Albert Y. S. Lam*

In this article, authors extend the optimal auction theory for application on a crowdsourcing platform, they first extend the results of the celebrated Myerson's optimal auction mechanism for a single-parameter bid to the case where the bid consists of the unit cost of work, the maximum amount of work one is willing to do, and the actual work completed. They show that a simple payment mechanism is sufficient to ensure a dominant strategy from the workers, it is advantageous for the worker to complete as much of the allocated work as possible, and propose and analyze a novel allocation mechanism of allowing the requester to balance between cost efficiency and equality in the allocation of work. Their main results are proved analytically and validated through simulations.

3. "A Learning-Embedded Attributed Petri Net to Optimize Student Learning in a Serious Game" by *Jing Liang, Ying Tang, Ryan Hare, Ben Wu, and Fei-Yue Wang*

This article develops a learning-embedded attribute Petri net (LAPN) model to represent game flow and student learning decision-making. The dynamics of learner behaviors in game are then addressed through the incorporation of learning mechanisms [i.e., reinforcement learning (RL) and random forest classification] into the Petri net model for knowledge reasoning and learning. Finally, the learning optimization algorithm is derived based on LAPN to aim to offer a faster and better solution to each problem-solving task in game. From the case study, the approach is found to be effective in decreasing game completion time. The benefit of the proposed model and algorithm is then demonstrated in the serious game (SG) gridlock.

4. "Community Sports Organization Development From a Social Network Evolution Perspective—Structures, Stages, and Stimulus" by *Jia Yu, Mengjun Ding, Qingran Wang, Weiqiang Sun, and Weisheng Hu*

In this article, the authors construct a dynamic social network from 43 747 teaming relations collected in a five-year period in a Community Sports Organization (CSO). They study the network structure and development stages of the CSO and investigate how external stimulus, in the form of online/offline and rewarded/unrewarded activities, may affect the development of the CSO. With a classification algorithm called stage classification based on community evolution events (SC-CEEs), they observe that the development of

the CSO can be robustly divided into stages. They further observe that the relationship between specific stimulus and community evolution varies across development stages. Then, they give practical guidance on the developmental priorities and response strategy of CSOs at different stages.

5. "Heterogeneous Network Representation Learning Approach for Ethereum Identity Identification" by *Yixian Wang, Zhaowei Liu, Jindong Xu, and Weiqing Yan*

In this article, the authors propose a heterogeneous network representation learning method to mine implicit information inside Ethereum transactions. Specifically, they construct an Ethereum transaction network by collecting transaction data from normal and phishing Ethereum accounts. Then, they propose a walk strategy that combines the timestamps and amounts of transactions to represent the characteristics, and then to mine the types of nodes and edges, they use a method of representation learning for attributed heterogeneous networks to map the transaction network to a low-dimensional space. The experimental results show that the heterogeneous network representation learning method outperforms existing algorithms for analysis of the Ethereum transaction dataset.

6. "Orthogonal-Moment-Based Attraction Measurement With Ocular Hints in Video-Watching Task" by *Minqiang Yang, Xiang Feng, Rong Ma, Xiongying Li, and Chengsheng Mao*

This article investigates the relationship between eye suggestion (i.e., pupil dilation signal and eye movement) and attractiveness level under the stimulation of emotional movie clips. This article also proposes two emotion classification feature extraction methods based on orthogonal moments and conducts comparative testing and ablation experiments. The results indicate that the method proposed in the article has better classification performance. In addition, the authors establish a new database of eye movement and pupil dilation signals, which can be used to study the attractiveness level under the stimulation of emotional movie clips.

7. "Graph Regularized Sparse Non-Negative Matrix Factorization for Clustering" by *Ping Deng, Tianrui Li, Hongjun Wang, Dexian Wang, Shi-Jinn Horng, and Rui Liu*

This article proposes a novel graph-regularized sparse non-negative matrix factorization (GSNMF) and its extension, namely, graph-regularized sparse nonnegative matrix trifactorization (GSNMTF). In order to obtain a cleaner data matrix to approximate a high-dimensional matrix, l1-norm is added to the low-dimensional matrix to increase feature sparsity and reduce the impact of noise. Based on this, corresponding reasoning and alternate iterative updating algorithms are proposed to solve the optimization problem. Finally, the experimental results show that the proposed model has good performance.

8. "Can Biases in Perceived Attitudes Explain Anti-Conformism?" by *Guillaume Deffuant, Omid Roozmand, Sylvie Huet, Kamilla Khamzina, Armelle Nugier, and Serge Guimond*

This article assumes that some biases about attitude perception can explain anticonformists. In order to study this hypothesis, the authors adopt a modeling method and design a computational model for perceptual bias, intention, and perceived group norms. The model calculates intentions based on reasoning action theory and assumes that there are some biases in the perception of others' attitudes based on social judgment theory. It studies the conditions for model parameter values, simulating and reproducing the characteristics observed in the study. The hypothesis is verified through the research results.

9. "Behavioral Model for Live Detection of Apps Based Attack" by *Misbah Shafi, Rakesh Kumar Jha, and Sanjeev Jain*

This article proposes an application-based attack modeling and detection method, which includes vulnerable end-user applications installed on smartphones. The detection model involves the proposed technology of application-based behavioral model analysis (ABMA) solutions to address attack models. ABMA is estimated using power, battery power, and data usage. A comparative analysis of application intrusion detection is conducted. The simulation results verify the effectiveness of the model.

10. "Analysis of Moral Judgment on Reddit" by *Nicholas Botzer, Shawn Gu, and Tim Weninger*

This work focuses on analyzing moral judgments rendered on social media by capturing the moral judgments that are passed in the subreddit /r/AmITheAsshole on Reddit. This article trains a classifier, uses a human annotator to verify the performance of this classifier, and uses it to investigate various website features surrounding moral judgment in the other ten subsections. The experimental results show that when a post contains positive moral values, users will vote more frequently to support the post.

11. "Output–Input Ratio Maximization for Online Social Networks: Algorithms and Analyses" by *Shengminjie Chen, Wenguo Yang, Yapu Zhang, and Suixiang Gao*

This article proposes a new problem, which considers the profit maximization problem from the perspective of the ratio of revenue to cost, called the output-to-input ratio maximization (OIRM). In order to solve the OIRM problem, the hill-climbing greedy algorithm, the threshold decrease algorithm, and the double greedy algorithm are designed for OIRM. Finally, through a large number of experiments to evaluate the performance of the algorithm, it is concluded that the threshold decrease algorithm is an effective algorithm to solve OIRM.

12. "The Effect of Facial Perception and Academic Performance on Social Centrality" by *Dongyu Zhang, Ciyuan Peng, Xiaojun Chang, and Feng Xia*

This article explores how facial expressions affect students' social centrality in various educational networks compared to actual academic performance. The authors first analyze the dual effects of facial perception and actual academic performance on the social centrality of students. Then, a novel network situation prediction framework based on face perception and actual academic performance is proposed. Finally, this study contributes to a comprehensive and in-depth understanding of social networks by analyzing social biases based on facial features.

13. "In Your Eyes: Modality Disentangling for Personality Analysis in Short Video" by *Xiangguo Sun, Bo Liu, Liya Ai, Danni Liu, Qing Meng, and Jiuxin Cao*

This article proposes a new video analysis framework for visual, auditory, and text neural network personality detection. In order to improve the sensitivity of the model to personality detection, three deep learning channels are proposed to learn modal features. The framework can not only extract each modal feature but also learn time-varying patterns via a temporal alignment network. A well-defined loss function is carefully designed, including prediction loss, difference loss of private features, and similarity loss of shared features. A large number of experiments on real-world datasets have demonstrated the advantages of our model.

14. "Agent-Based Model of Mass Campus Shooting: Comparing Hiding and Moving of Civilians" by *Peng Lu, Dianhan Chen, Yan Li, Xiaoyong Wang, and Shanping Yu*

Considering the dichotomous characteristics of the people shot and the impact of the refuge area on the crowd movement in the shooting incident, this article uses an agent-based model to restore the real target case. Based on the optimal solution, the authors explore the effects of the number and location of building exits, the scale of the shooter's attack, and the time to find shelter on the shooting results. At the same time, they propose the critical point of hiding probability.

15. "Time-Aware Attention-Based Gated Network for Credit Card Fraud Detection by Extracting Transactional Behaviors" by *Yu Xie, Guanjun Liu, Chungang Yan, Changjun Jiang, and MengChu Zhou*

This article proposes a novel model to extract the transactional behaviors of users and learn new transactional behavioral representations for credit card fraud detection. Considering the characteristics of transactional behaviors, two time-aware gates are designed in a recurrent neural net unit to learn long- and short-term transactional habits of users, respectively, and to capture behavioral changes of users caused by different time intervals between their consecutive transactions. In addition, an interaction module is designed to enable the model to learn more comprehensive and rational representations.

16. "Forecasting COVID-19 Outbreak Through Fusion of Internet Search, Social Media, and Air Quality Data: A Retrospective Study in Indian Context" by *Sankhadeep Chatterjee, Kushankur Ghosh, Arghasree Banerjee, and Soumen Banerjee*

This article proposes a machine learning augmented technique to predict the coronavirus disease (COVID-19) outbreak in India by combining Internet search trends along with social media data retrieved from Twitter. A comprehensive list of suitable search words has been used to select a large collection of Tweets, and the Internet search trends of the same keywords have been fetched. A lag correlation analysis is conducted to find the number of days, and both shallow and deep learning methods are engaged to predict the number of COVID-19 cases.

17. "Unity in Diversity: Multilabel Emoji Identification in Tweets" by *Gopendra Vikram Singh, Mauajama Firdaus, Asif Ekbal, and Pushpak Bhattacharyya*

This article proposes a multilabel emoji prediction system that predicts the appropriate emoji for a given tweet by using different state-of-the-art baselines. They create a large-scale multilabel emoji dataset named Mu-Emoji that comprises of more than 0.6 million tweets. They employ a graph attention network along with bidirectional encoder representations from the transformer encoder for the accurate prediction of emojis. The qualitative and quantitative analyses show that the multilabel emoji prediction baselines perform well.

18. "PIWS: Private Intersection Weighted Sum Protocol for Privacy-Preserving Score-Based Voting With Perfect Ballot Secrecy" by *Siqi Lu, Zhaoxuan Li, Xuyang Miao, Qingdi Han, and Jianhua Zheng*

This article proposes private intersection weighted sum (PIWS), a scalable, fair, and privacy-preserving intersection-weighted sum protocol. The PIWS protocol can privately calculate the intersection of identity index sets maintained by each participant and can privately calculate the weighted sum of the data associated with the identity indexes of the intersection set. The protocol not only achieves the function of an optional weighted voting protocol but also is relatively lightweight and efficient.

19. "Emergence of Social Norms in Metanorms Game With High-Order Interaction Topology" by *Zhen Wang, Ruodan Li, Xing Jin, and Hong Ding*

This article introduces hypergraphs in Axelrod's metanorm game model, which are more suitable to reflect group interaction to model metanorm games. The establishment efforts of norms are examined on hypergraphs (uniform random hypergraph (URH), hyperdegree-heterogeneous random hypergraphs, and real-world hypergraphs). The difference in group sizes affects the norms emergence of agents and realizes the establishment of social norms on URH. The authors propose the dynamic relevance boldness and vengefulness learning (BV-learning) algorithm to overcome the aforementioned problems. Finally, in order to illustrate the universality of the dynamic relevance BV-learning clearly, they study the norms establishment on real-world hypergraphs and compare it with BV-learning.

20. "Opinion Dynamics of Social Networks With Intermittent-Influence Leaders" by *Zijie Zhao, Lei Shi, Tong Li, Jinliang Shao, and Yuhua Cheng*

This article constructs a leader–follower architecture by introducing intermittent-influence opinion leaders to the DeGroot model and analyzes the influence of this type of leader on the evolution of opinions. The leaders in this article can only convey its opinion by broadcasting at some intermittent moments. They analyze the relationship between the leaders' broadcast moments and the consensus opinion of followers. Then, they describe the connotation of assimilation and calculate the minimum number of broadcasts required for the leaders to assimilate the follower's opinions.

21. "Exploring Public Sentiment During COVID-19: A Cross Country Analysis" by *Shuo Yu, Sihan He, Zhen Cai, Ivan Lee, Mehdi Naseriparsa, and Feng Xia*

In this work, the authors investigate social media posts on Twitter and Weibo in order to explore the difference in reactions across various countries, and to understand national differences. They employ natural language processing (NLP) methods and linguistic inquiry and word count (LIWC) tools to process six languages in different countries, including the United States, Germany, France, Italy, the United Kingdom, and China, and provide a comprehensive analysis of public reaction differences from the emotional perspective. The findings verify that the reactions vary noticeably among various countries for some policies to make appropriate policy decisions.

22. "Predicting Politeness Variations in Goal-Oriented Conversations" by *Kshitij Mishra, Mauajama Firdaus, and Asif Ekbal*

In this article, the authors create politeness-annotated conversational data (PACD) utilizing Microsoft dialog challenge and dialog state tracking challenge 1 (DSTC1) datasets to identify politeness in goal-oriented dialog systems. They employ a hierarchical transformer network that effectively captures the contextual information (i.e., previous utterances) and current input for predicting the politeness in a given utterance of a dialogue. The empirical results demonstrate that the proposed approach outperforms all the defined baselines. Furthermore, through in- and cross-domain experiments, they show the necessity of a PACD to mitigate acts such as rude requests or insults for both socially interactive and task-oriented dialog systems.

23. "A Hybridized Stochastic SIR-Vasiček Model in Evaluating a Pandemic Emergency Financing Facility" by *Weili Fan and Rogemar Mamon*

This article examines the design and valuation of a pandemic emergency financing facility (PEFF) akin to a catastrophe (CAT) bond. At time t, $S(t)$, $I(t)$, and $R(t)$ satisfy a system of interacting stochastic differential equations in our extended framework. The payout is triggered when the number of infected people exceeds a predetermined threshold. A CAT-bond pricing setup is developed with the Vasiček-based financial risk factor correlated with the SIR dynamics for the PEFF valuation. The analyses reveal that the SIR's disease transmission and recovery rates, as well as the interest rates' mean-reverting level, have a substantial effect on the bond price.

24. "CTL-DIFF: Control Information Diffusion in Social Network by Structure Optimization" by *Jinyin Chen, Xiaodong Xu, Lihong Chen, Zhongyuan Ruan, Zhaoyan Ming, and Yi Liu*

In this article, they propose a diffusion control method based on gradient information to generate an optimized network structure, namely, ConTroL information DIFFusion (CTL-DIFF). It first selects the edges with the largest absolute gradient based on the prediction model to optimize the original network's structure. It then employs several prediction methods to verify whether the target user's social action status is controlled. After comparing with four baselines on three datasets. The experimental results show that CTL-DIFF can effectively control information diffusion in the global social network by identifying and controlling the most influential users.

25. "Mental Health Disorder Identification From Motivational Conversations" by *Tulika Saha, Saichethan Miriyala Reddy, Sriparna Saha, and Pushpak Bhattacharyya*

Distinguishing and identifying various mental disorders is a challenging task. This work presents a hierarchical attention-based deep neural network classifier for modeling conversations to detect different mental disorders as the dialog progresses. The proposed network is equipped with lexicon-based sentiment scores in order to prioritize certain contributing features. The proposed model attained an accuracy of 83.91% and outperformed several strong baselines.

26. "Performance Comparison of Transformer-Based Models on Twitter Health Mention Classification" by *Pervaiz Iqbal Khan, Imran Razzak, Andreas Dengel, and Sheraz Ahmed*

Transformer-based methods are better at capturing the meaning of a word based on its surrounding words compared to traditional methods in health mention classification. This work with nine widely used transformer methods and compare their performance on the personal health mention classification of tweet data. They analyze the impact of model size on the classification task and provide a brief interpretation of the classification decision made by the best-performing classifier. The experimental results show that RoBERTa outperforms all other models by achieving an $F1$ score of 93%, while two other models perform similarly by achieving an $F1$ score of 92.5%.

27. "COVID-19 Mortality Risk Factors Using Survival Analysis: A Case Study of Jakarta, Indonesia" by *Bahrul I. Nasution, Yudhistira Nugraha, Nanda L. Prasetya, Muhammad E. Aminanto, Andi Sulasikin, Juan I. Kanggrawan, and Alex L. Suherman*

There are limited studies regarding the factors of COVID-19 mortality in Indonesia with a more balanced dataset. The previous studies only implemented logistic regression, sensitive to the imbalanced dataset. This study aims to identify the risk factors for COVID-19 mortality in Indonesia using a survival analysis approach using Jakarta as a case study. They use the piecewise exponential model (PEM) to overcome the time-dependent problem in CPH. The findings show that COVID-19 mortality does not differ between genders, but the elderly are 3.5 times more likely to die from the virus. Dyspnea, malaise, and pneumonia are the primary symptoms associated with COVID-19 mortality.

28. "Social Media in Emerging Economies: A Cross-Cultural Comparison" by *Thomas F. Stafford and Bao Q. Duong*

In the highly limited context of a grounded theory study of specific Vietnamese and Singaporean users of social media and social commerce platforms, the article compared instances of both an emerging market and an advanced market in the hemispheric region. The authors finally found that the Vietnamese culture of social media use orients around economic productivity, social commerce, and leveraging social connections

to largely generate consumer-to-consumer commerce, but the Singaporean culture of social media revolves around Influence Commerce with the government's support.

29. "Supervised Attention Network for Arbitrary-Shaped Text Detection in Edge-Fainted Noisy Scene Images" by *Aishwarya Soni, Tanima Dutta, Nitika Nigam, Deepali Verma, and Hari Prabhat Gupta*

For detecting arbitrary-shaped text instances in noisy scene images with faint text edges, the authors propose a semantic edge supervised spatial-channel attention network (SESANet). SESANet is efficient, precise, and fast in nature. In order to precisely localize the text masks in scene images with poor contrast and illumination, the network learns multiscale supervised edge semantic, pixel-wise spatial structure information, and interchannel dependencies. The experimental results show superior performance with regard to recall on the publicly available benchmark datasets.

30. "A Matrix Factorization Recommendation System-Based Local Differential Privacy for Protecting Users' Sensitive Data" by *Xiaoyao Zheng, Manping Guan, Xianmin Jia, Liangmin Guo, and Yonglong Luo*

In this article, the authors propose a novel matrix factorization (MF) algorithm to achieve a balance between recommendation performance and privacy-preserving of recommendation system (RS). Then, based on the above algorithm, this article proposes an MF RS for preserving user privacy by using local differential privacy technology. After that, Laplace noise is added to the users' sensitive data before it is sent to the data aggregator. Finally, based on the disturbed data, the rating prediction is realized following the MF algorithm.

31. "Tabular Learning-Based Traffic Event Prediction for Intelligent Social Transportation System" by *Chen Sun, Shen Li, Dongpu Cao, Fei-Yue Wang, and Amir Khajepour*

The authors propose a framework to integrate the social traffic data and use the TabNet model to facilitate the representation learning task in traffic event prediction. The TabNet model inherits the advantages of the decision tree-based approach and the advantages of a deep neural network. The experimental results show the TabNet model performs better in traffic event prediction for its instance interpretation capability and unsupervised pretraining. The study has great practical significance for regulating traffic planning and other fields.

32. "Hashtag-Based Tweet Expansion for Improved Topic Modeling" by *Durgesh Kumar, Loitongbam Gyanendro Singh, and Sanasam Ranbir Singh*

This study uses internal vocabulary (hashtags) to counter under-specificity and sparsity in tweets and proposes a unified framework for hashtag-based tweet expansion. First, the authors evaluated the importance of hashtags in Latent Dirichlet Allocation (LDA) performance and the result shows hashtags are an important feature for finding topics. Furthermore, to expand tweets with semantically related hashtags, they get the tweet representation using textual content successfully. LDA performance after expanding tweets with the proposed expansion approaches improves significantly compared to raw tweet and hashtag pooling-based tweets expansion.

33. "The Impact of the Variability of Patient Flow and Service Time on the Efficiency of Large-Scale Outpatient Systems" by *Chengye Zou, Junwei Wang, and Yao Cheng*

In order to find the effect of patient flow variability and service time variability on the efficiency of a multistage outpatient system, this article built a discrete-event simulation model and propose a novel method to generate the transition matrix based on patient flow variability. The experimental results show that patient flow variability has impacts on patient satisfaction, resource utilization, and system efficiency. Service time variability has impacts on resource utilization, patient satisfaction, and system efficiency with patient flow variability.

34. "Pareto Improvement: A GRA Perspective" by *Haibin Zhu*

Pareto optimality refers to a benefit distribution that improvement cannot be achieved if the benefit of at least one person is not reduced. In this article, the author contributes a way to examine Pareto improvements and optimization from the viewpoint of role assignments, using different role assignment models to simulate resource/benefit allocations in a society. The author deduced some interesting findings, such as Pareto improvements do not exist when resources are insufficient and are feasible only when the number of provisions is larger than that of requests.

35. "Comparing the Impact of Social Media Regulations on News Consumption" by *Gabriele Etta, Matteo Cinelli, Alessandro Galeazzi, Carlo Michele Valensise, Walter Quattrociocchi, and Mauro Conti*

In this article, the authors analyze two social media that enforced opposite moderation methods, Twitter and Gab, to assess the interplay between news consumption and content regulation concerning COVID-19. Their results show that the presence of moderation pursued by Twitter produces a significant reduction in questionable content. Conversely, the lack of clear regulation on Gab results in the tendency of the user to engage with both types of content, showing a slight preference for the questionable ones which may account for a dissing/endorsement behavior. They conclude that content policies cover an important role against the circulation of harmful content, especially in the context of the COVID-19 pandemic.

36. "Semantic Tradeoff for Heterogeneous Graph Embedding" by *Yunfei He, Dengcheng Yan, Yiwen Zhang, Qiang He, and Yun Yang*

In this article, the authors propose semantic tradeoff heterogeneous graph embedding (STHGE) by first introducing the Hilbert–Schmidt independence criterion (HSIC) as a restriction. The main idea of STHGE is to regard semantic tradeoff as an independence tradeoff (or correlation) between different meta-path spaces. To evaluate the effectiveness and feasibility of STHGE, They conducted experiments using three publicly available heterogeneous graph (HG) datasets. The experimental results demonstrated the feasibility and effectiveness of STHGE. Moreover, they not only demonstrated that semantic tradeoffs exist in these datasets but also verified that STHGE is robust. Furthermore, STHGE provides a novel way to gain insight into the relationships between HG meta-paths.

37. "Walrasian Equilibrium-Based Pricing Mechanism for Health-Data Crowdsensing Under Information Asymmetry" by *Xinxin Guo, Nan Kong, and Haiyan Wang*

In this article, the authors model the strategic interactions between health-data requesters and collectors using a bilevel optimization model. They propose a Walrasian equilibrium-based pricing mechanism to coordinate the interest conflicts between health-data requesters and collectors. They prove that the optimal task price is equal to the marginal utility generated by the collector's health data. To avoid obtaining the collector's private information, a distributed iterative algorithm is then designed to obtain the optimal task pricing strategy. Based on the proposed pricing mechanism, on the one hand, the requester and collectors can maximize their payoffs simultaneously; on the other hand, the proposed pricing mechanism can be easy to implement while ensuring truthfulness and fairness.

38. "PIANO: Influence Maximization Meets Deep Reinforcement Learning" by *Hui Li, Mengting Xu, Sourav S. Bhowmick, Joty Shafiq Rayhan, Changsheng Sun, and Jiangtao Cui*

In this article, the authors present an orthogonal and novel paradigm to address the influence maximization (IM) problem by leveraging deep reinforcement learning (RL) to estimate the expected influence. In particular, they present a novel framework called deeP reInforcement leArning-based iNfluence maximizatiOn (PIANO) that incorporates network embedding and RL techniques to address this problem. In order to make it practical, the authors further present PIANO-E and PIANO at $<d>$, both of which can be applied directly to answer IM without training the model from scratch.

39. "FaiRIR: Mitigating Exposure Bias From Related Item Recommendations in Two-Sided Platforms" by *Abhisek Dash, Abhijnan Chakraborty, Saptarshi Ghosh, Animesh Mukherjee, and Krishna P. Gummadi*

The authors introduce multiple flexible interventions [fair related item recommendation (FaiRIR)] in the RIR pipeline. they instantiate these mechanisms with two well-known algorithms for constructing RIRs—rating singular value decomposition (SVD) and item2vec—and show on real-world data that the proposed mechanisms allow for a fine-grained control on exposure distribution, often at a small or no cost in terms of recommendation quality, measured in terms of relatedness and user satisfaction. The proposed algorithms are applicable to any other domain, including job recommendation sites and others. In this work, they considered RIRs that only use relatedness with respect to one source item.

40. "Kernel-Induced Possibilistic Fuzzy Associate Background Subtraction for Video Scene" by *Badri Narayan Subudhi, Manoj Kumar Panda, T. Veerakumar, Vinit Jakhetiya, and S. Esakkirajan*

The authors propound an idea of a kernel-induced possibilistic fuzzy associated background subtraction (BGS) scheme for local change detection from a fixed camera-captured sequence. The proposed scheme follows two stages: background training and foreground segmentation. In the background construction stage, each pixel is modeled using a possibilistic fuzzy cost function in kernel-induced space. The performance of the

proposed scheme is tested on three benchmark databases. The effectiveness of the proposed scheme is evaluated on different performance evaluation measures: precision, recall, F-measure, and average similarity.

41. "Few Sample Generation of Amount in Figures for Financial Multi-Bill Scene Based on GAN" by *Zhi-Ri Tang, Qi-Qi Chen, Zhao-Hui Sun, Pengwen Xiong, Bao-Hua Zhang, Lu Jiang, and Edmond Q. Wu*

Aiming at the problem of insufficient training data in multi-bill scenes and the low accuracy of the detection model, this article proposes a new generative adversarial network (GAN) to generate new samples and to expand the bill dataset. In the proposed network, a residual block is adopted as the basic structure of the generator and the discriminator, and the self-attention mechanism is also utilized to improve the generation performance, which is then adopted to train a framework for recognition of the bill amount.

42. "An Evaluation of Technology Adoption During Remote Teaching and Learning at Tertiary Institution by Gender" by *Kehinde Aruleba, Nobert Jere, and Obert Matarirano*

This study examined the impact of the rapid transition from face-to-face learning to remote learning (RL) for students. The study employed a quantitative census sampling method as an electronic questionnaire administered to 243 respondents. In designing the questionnaire, ten items from the Technology Readiness Index (TRI 2.0) were adapted to determine the technology readiness of respondents. Furthermore, descriptive statistics were used to establish the common information and communications technology (ICT) tools, platforms, and teaching approaches adopted by lecturers and estimate the technological readiness of students.

43. "Modeling and Analysis of Competitive Behavior in Social Systems" by *Suibing Li, Long Jin, and Shuai Li*

In this article a new competition model is developed, which aims to describe the competitive behavior in social systems. They define various components of social networks as some parameters and use changes in parameters to describe the dynamic changes in opinions. This article proves the stability and convergence of the constructed competition model in theory. A series of simulation experiments are conducted to simulate competitive activities in real life, and application scenarios suitable for the model are provided in this work.

44. "Measuring Network Polarization and Political Sectarianism During the 2020 Pandemic" by *Carlos Kamienski, Claudio Luis de Camargo Penteado, Denise Goya, Rafaela V. Rocha, Lucas Mazim de Souza, Daniel Vitor Beraldo di Gênova, Diogo Fornaziero Segura Ramos, Fabrício Olivetti de França, Flávio Horita, and Carlos da Silva dos Santos*

This article proposes an approach for measuring network polarization and political sectarianism in Twitter based on user interaction networks. This network polarization and the measurement of political sectarianism are computed by the Bayesian probability. The authors collected Twitter data from 33 conflicted political events in Brazil during 2020, strongly influenced by the COVID-19 pandemic. The results reveal that the approach based on user interaction networks leads to an increasing understanding of polarized conflicts in Twitter

and a small number of polarizers is enough to represent the polarization and sectarianism of Twitter events.

45. "Semantic Learning for Facial Action Unit Detection" by *Xuehan Wang, C. L. Philip Chen, Haozhang Yuan, and Tong Zhang*

This article proposes an action unit (AU) detection architecture based on semantic embedding for image transformers (SEiTs). The authors proposed that SEiT can inherently learn morphological features from facial images. The pretraining task aims to learn the semantic representation of corresponding facial components. In addition, the learned semantic embeddings are fed to the transformer blocks to achieve global interaction between semantic elements. SEiT naturally integrates facial morphology information and global interactive features, making it very suitable for AU detection. The experimental results demonstrate the superiority of SEiT.

46. "Crisis Assessment Oriented Influence Maximization in Social Networks" by *Weinan Niu, Wenan Tan, Wei Jia, Lu Zhao, and Na Xie*

This article introduces a crisis assessment-oriented and topic-based influence maximization problem (TIM-CA). This problem models the influence maximization (IM) problem by considering the crisis assessment (CA) and the topics of users. In addition, the authors propose a maximum influence arborescence model-based algorithm for TIM-CA, namely, MIA-TIM-CA. It adopts a new influence propagation evaluation function that combines topic relevance, crisis level, and maximum influence tree. For crisis degree calculation, it considers the crisis degree of each node based on itself, topology, and topic. The experimental results indicate that MIA-TIM-CA is superior to other state-of-the-art IM methods.

47. "Internet Financial Fraud Detection Based on Graph Learning" by *Ranran Li, Zhaowei Liu, Yuanqing Ma, Dong Yang, and Shuaijie Sun*

In this article, the authors proposed a graph-learning algorithm for Internet financial fraud detection. The algorithm can learn transaction amount features and topological features in a financial transaction network graph and represent them as low-dimensional dense vectors. Intelligent and efficient classification and prediction are allowed by training classifier models. In future work, the method will be improved and implemented with spatio-temporal properties. By doing this, the method can effectively learn the features of newly generated vertices in a dynamic network graph to achieve better financial fraud detection.

48. "Positive Evaluation Maximization in Social Networks: Model and Algorithm" by *Shengminjie Chen, Wenguo Yang, and Suixiang Gao*

Great attention is paid to influence maximization (IM), which is a classical and heated issue in online social networks. In this article, the authors propose a novel problem called positive evaluation maximization (PEM) and construct a novel propagation model named MFP-independent cascade (MFP-IC). They prove the PEM problem can be rewritten as a difference between submodular (DS) decomposition. In order to solve the problem, the authors design two novel parametric-conditioned greedy (PCG) algorithms based on DS

decomposition. To speed up algorithms, they improve PCG to fast PCG (FPCG) using a sampling technique.

49. "COVID-19 Pandemic Severity Criterion Based on the Number of Deaths and the Uneven Distribution of These" by *Antoni Wilinski, M. K. Arti, and Lukasz Kupracz*

The aim of this article is to define a criterion related to the number of Coronavirus disease 2019 (COVID-19) deaths in different countries to compare them themselves in two aspects. By considering the population numbers in these units, the Lorentz curve is prepared, and the Gini coefficient is calculated for the entire world and for individual countries such as the USA, India, Brazil, Poland, and the Balkan and Eastern European countries with the highest number of deaths per million inhabitants in the world. Moreover, an attempt was made to present the universal mortality rate in a given country in the form of a bicriterion combining the Gini index and the number of deaths per million inhabitants achieved.

II. GROWING SOFTWARE: OBJECTIVE, METHODOLOGY, AND TECHNOLOGY

With the continuous advances of the digital transformation, software is increasingly recognized for its infrastructural role in our modern society. In fact, "our civilization runs on software," as Bjarne Stroustrup put it. However, software construction and maintenance are getting inevitably complex with the rapid growth of software functionalities as requested by our expanding industry and business requirements. This situation has started to overwhelm the capability limits of even professional software developers. Furthermore, new computing technologies, such as cloud computing, big data, the Internet of Things, and deep learning, are constantly increasing the need for addressing the everlasting changes in software's running environments, dependent resources, and functional requirements. Such trends, deviating from the classical assumptions of software engineering methodologies and technologies on the fixedness or slow evolution of software environments, resources, and requirements, have presented nonnegligible challenges to software researchers and practitioners.

A. Making Software Survive and Grow

The challenges have attracted much attention from the academia and governments targeting at future social advances built on software infrastructures. For example, the defense advanced research projects agency (DARPA) of the USA proposed the building resource adaptive software systems (BRASSs) program project, aiming "to realize foundational advances in the design and implementation of long-lived, survivable and complex software systems that are robust to changes in the physical and logical resources provided by their ecosystem" [1], and its department of defense (DoD) recently approved the software modernization strategy, focusing on "the ability to securely and rapidly deliver resilient software capability" and requesting "the ability to evolve faster and be more adaptable" [2]. The government of China also organized research efforts on the methodologies and technologies for continuously evolving software, aiming to make the software

survive longer and adapt to changes by itself. Then, a foundational question naturally arises: *can software possibly grow by itself?*

By *growing*, the authors mean that software should be capable of continuously evolving from its initial version without or with little human aid, in order to cope with the everlasting changes in its running environments, dependent resources, and functional requirements, by continuously adding functionalities to, and fixing bugs in, its original code. Such a continuously evolving approach differs from its traditional counterpart that assumes the completeness of all software requirements and their slow evolution. As such, *growing software* must call for novel changes in its theoretical foundation and development paradigm. The authors have made some preliminary explorations, and try to summarize our thoughts and results into two aspects, namely, *methodological framework* and *realization routes* [3], as follows.

1) Regarding the *methodological framework*, the authors consider that growing software should own three key characteristics, namely, *three-folded architecture model*, explicit runtime adaptation mechanism, and continuously evolving development lifecycle. The three-folded architecture model connects represented users (for realizing their business values), sensible environments (for understanding the environmental dynamics), and evolvable systems (for coordinating the software evolution) together, making the complete modeling explicit about the software and its environments and users. The explicit runtime adaptation mechanism implies a four-step loop in the software execution, which includes environment-sensing, decision-making, dynamic update, and continued execution, enabling the software's iterative adaptation capability beyond its original business logic. Finally, the continuously evolving development lifecycle requests a progressively optimizing procedure, as driven by users and developers, on the software's functionality addition and bug fixing, in a continuous and never-ending manner.

2) Regarding the realization routes, the authors consider that the characteristics of growing software imply three important observations. First, growing software can be realized by following the three characteristics when developing the software from the very beginning, e.g., modeling users, environments, and systems explicitly, and designing the four-step loop and progressively optimizing the procedures together with the target business logic. Second, regarding legacy software, transforming it into a growing one can be realized by extracting internal adaptive potentials from the legacy software and exploiting them for enhanced adaptability at runtime, e.g., one possibility is mining and extracting the resources the software depends on and their relationships with the software's functionalities and performance, and later configuring the software and guarding its quality by observing the invariants derived from the extracted resources and relationships. Third, growing software can also be realized in a data-driven way by making the software development process intelligent and automated with minimized human aids, e.g., mining the construction and evolution rules and strategies from past software development histories, and supporting automated functionality addition and bug fixing in later similar software evolution tasks.

B. Enabling Technologies for Growing Software

Along with the study of growing software, the authors consider that this new paradigm of software would call for a wide range of enabling technologies. Here, the authors briefly review a few published technologies for example, which have advanced the research progress in this field as follows.

1) *On Correctly Understanding Environments:* GEAS [4] can automatically analyze hidden interferences among context changes, and conducts constraint checking with adaptive grouping based on such analysis. Such adaptively grouped constraint checking can greatly improve the efficiency (up to $28.6\times$) with a soundness guarantee for the consistency of application contexts, making the applications understand dynamic environments correctly and efficiently.

2) *On Dynamically Updating Software:* PASTA [5] can automatically synthesize across-version object state transformers from multi-version code, with a greatly improved success rate ($7.5\times$). This synthesis automation makes it possible to update the software dynamically from an old version to its new version at runtime, without the need for software shutdown and restart.

3) *On Automatically Configuring Software:* SafeTune [6] can mine and extract the relationships (8000 from 16 applications) between software configurations and intentions, and automatically identify performance-related configuration issues (tens) with greatly improved efficiency (hours down to seconds). This enables applications to know how to reconfigure themselves for better performance with a safety guarantee.

4) *On Soundly Fixing Bugs:* Recoder [7] can automatically generate patches and fix bugs for software failures. It is based on deep neural network representations trained on massive code and has largely expanded the auto-bug-fixing application by its greatly improved success rate ($2.4\times$).

Image an application scenario in which the authors develop and deploy a software system that can support auto-pilot features for modern vehicles. Such a software system is typically very complex, in the sense that they combine software and hardware, as well as code logics and machine learning models. Such a mixture clearly complicates the software design, development, and maintenance. Now consider how the paradigm of growing software and its enabling technologies help as follows.

1) First, the software system can be initially developed with its built-in sensing, adaptation, and evolution support. Then, the system's sensing part can transparently conduct efficient context consistency checking (e.g., by GEAS [4]) to ensure that the system correctly understands its running environment (i.e., validating the

massive data received from the vehicles' sensors), and knows when it should react to environmental changes or call for reconfiguration or evolution.

2) Second, when a reconfiguration is needed for coping with environmental changes (e.g., changes in its dependent resources), the system can conduct performance tuning (e.g., by SafeTune [6]) to find the new balance point (i.e., supporting stable driving even when its available resources are reduced), while not breaking the invariants implied by the underlying resource-intention-configuration relationships (i.e., still guaranteeing the safe-driving principle).

3) Third, when an evolution is needed for replacing the original code for implementing new functionalities, the system can conduct dynamic software update (e.g., by PASTA [5]) to transform the system from its old version to the new version without shutdown and restart. This feature is highly desirable (e.g., the vehicle can thus upgrade its system version anytime, even during driving).

4) Finally, the system can contain small bugs here and there, and debugging such a complex system is painful. Automated tools for bug localization and fixing can help much (e.g., by Recoder [7]), particularly useful for finding hidden bugs anywhere (e.g., in code logics or trained models).

C. Conclusion

With the discussed methodological and technical advances and examples, the authors observe that growing software has quite potential, and its idea and application can grow rapidly by itself. The study of growing software is attracting more software researchers and practitioners, and producing more useful results that can effectively support smart adaptation and continuous evolution of advanced software systems. Besides the current achievements, growing software still needs further efforts, for example, regarding the aforementioned auto-pilot software development and maintenance scenario, how to distinguish normal sensory inputs from adversarial inputs, how to optimally decide when to call for a system reconfiguration or evolution, how to guarantee bug-fixing would not introduce unexpected bugs, and so on.

JIAN LU
State Key Laboratory of Novel Software Technology
Nanjing University
Nanjing 210023, China
e-mail: lj@nju.edu.cn

CHANG XU
State Key Laboratory of Novel Software Technology
Nanjing University
Nanjing 210023, China
e-mail: changxu@nju.edu.cn

XIAOXING MA
State Key Laboratory of Novel Software Technology
Nanjing University
Nanjing 210023, China
e-mail: xxm@nju.edu.cn

BIN HU, *Editor-in-Chief*
School of Medical Technology
Beijing Institute of Technology
Beijing 100081, China
e-mail: tcss.ieee@gmail.com

REFERENCES

[1] *Building Resource Adaptive Software Systems (BRASS)*. (2023). [Online]. Available: https://www.darpa.mil/program/building-resource-adaptive-software-systems

[2] *DoD Software Modernization Strategy Approved*. (2023). [Online]. Available: https://www.defense.gov/news/releases/release/article/2924187/dod-software-modernization-strategy-approved

[3] C. Xu, Y. Qin, P. Yu, C. Cao, and J. Lu, "Theories and techniques for growing software: Paradigm and beyond," *Scientia Sinica Informationis*, vol. 50, no. 11, pp. 1595–1611, 2020.

[4] H. Wang, C. Xu, B. Guo, X. Ma, and J. Lu, "Generic adaptive scheduling for efficient context inconsistency detection," *IEEE Trans. Softw. Eng.*, vol. 47, no. 3, pp. 464–497, Mar. 2021.

[5] Z. Zhao, Y. Jiang, C. Xu, T. Gu, and X. Ma, "Synthesizing object state transformers for dynamic software updates," in *Proc. IEEE/ACM 43rd Int. Conf. Softw. Eng. (ICSE)*, Madrid, Spain, May 2021, pp. 1111–1122.

[6] H. He et al., "Multi-intention-aware configuration selection for performance tuning," in *Proc. 44th Int. Conf. Softw. Eng.*, Pittsburgh, PA, USA, May 2022, pp. 1431–1442.

[7] Q. Zhu et al., "A syntax-guided edit decoder for neural program repair," in *Proc. 29th ACM Joint Meeting Eur. Softw. Eng. Conf. Symp. Found. Softw. Eng.*, Athens, Greece, Aug. 2021, pp. 341–353.

Jian Lu (Member, IEEE) received the Ph.D. degree in computer science and technology from Nanjing University, Nanjing, China, in 1988.

He is currently a Full Professor with the Department of Computer Science and Technology and the Director of the State Key Laboratory for Novel Software Technology, Nanjing University. He is also an Academician of the Chinese Academy of Sciences, Beijing, China. His research interests include software methodologies, automated software engineering, formal methods, and middleware systems.

Chang Xu (Senior Member, IEEE) received the Ph.D. degree in computer science and engineering from The Hong Kong University of Science and Technology, Hong Kong, China, in 2008.

He is currently a Full Professor with the State Key Laboratory for Novel Software Technology and the Department of Computer Science and Technology, Nanjing University, Nanjing, China. His research interests include big data software engineering, intelligent software testing and analysis, and adaptive and autonomous software systems.

Dr. Xu is a Senior Member of the ACM. He participates actively in programs and organizing committees of major international software engineering conferences.

Xiaoxing Ma received the Ph.D. degree in computer science from Nanjing University, Nanjing, China, in 2003.

He is currently a Full Professor with the State Key Laboratory for Novel Software Technology and the Department of Computer Science and Technology, Nanjing University. His current research interests include self-adaptive software systems, neuro-symbolic software systems, and software engineering for socio-cyber-physical systems.

Bin Hu (Fellow, IEEE) received the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Science, Beijing, China, in 1998.

He is a (Full) Professor and the Dean of the School of Medical Technology, Beijing Institute of Technology, Beijing. He has (co)authored more than 400 publications (10k+ citations, H-index: 51). He is a Principal Investigator for large grants such as the National Transformative Technology "Early Recognition and Intervention Technology of Mental Disorders Based on Psychophysiological Multimodal Information," which have greatly promoted the development of objective, quantitative diagnosis, and non-drug interventions for mental disorders.

Dr. Hu is a fellow of the IET. He is a National Distinguished Expert, the Chief Scientist of 973, and the National Advanced Worker in 2020. He is a member of the Steering Council of the ACM China Council. He is a member of the Computer Science Teaching and Steering Committee as well as the Science and Technology Committee. His awards include the 2014 China Overseas Innovation Talent Award, the 2016 Chinese Ministry of Education Technology Invention Award, the 2018 Chinese National Technology Invention Award, and the 2019 WIPO-CNIPA Award for the Chinese Outstanding Patented Invention. He is the Vice-Chair of the China Committee of the International Society for Social Neuroscience. He is also the TC Co-Chair of computational psychophysiology and cognitive computing in the IEEE Systems, Man, and Cybernetics Society (SMC). He serves as the Editor-in-Chief for IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS.