Resumes

from PhD Scientific Seminar

Department Computer Systems and Technologies

Faculty Eletronics and Automation

Technical University Sofia branch Plovdiv,

08 February 2019

Contents

Editor: Nikolay Kakanakov

Milena Angelova is currently a Ph.D student at the Department of Computer Systems and Technologies of the Technical University of Sofia – branch Plovdiv. She received a B.Sc. degree in Automation and Control Systems in 2013 and an M.Sc. degree in Computer System and Technologies in 2015 from the same university. Her Ph.D. thesis supervisors are Prof. Veselka Boeva from the Technical University of Sofia – branch Plovdiv and Elena Tsiporkova from Sirris (Belgium). The subject of Milena's Ph.D. thesis is "Intelligent approaches to profiling, analyzing and locating expertise from publicly available sources." Milena also is an assistant professor at the Department of Computer Systems and Technologies. Her research interests are focused on intelligent information access. This entails developing of retrieval techniques that support humans in dealing with massive volumes of data. She addresses her Ph. D. work on proposing efficient clustering techniques for incrementally modeling a massive volume of incoming data that making changes to the stored data frequently. Her primary research interests and activities are in the field of Information Retrieval Systems, Machine Learning, Semantic Web and Recommender Systems.

A Split-Merge Evolutionary Clustering Algorithm[1]

MILENA ANGELOVA

Computer Systems and Technologies Department
Technical University of Sofia-branch Plovdiv
Tsanko Dyustabanov 25, 4000 Plovdiv, Bulgaria
mangelova@tu-plovdiv.bg

Most of the organizations are continually dealing with data that comes from their employees, customers, potential candidates or their external sources. Applications as expertise retrieval and healthcare decision support systems adapt their databases with available information periodically. Typically, the stored data is partitioned in disjoint clusters. The existing techniques as incremental and data stream algorithms present some challenges; incremental algorithms process one data element at a time and then re-cluster the data while to be merged in a single clustering solution; data stream algorithms lead to scalability issues and memory limitations. Practically, it is becoming impractical to re-cluster the massive amount of data whenever a new portion of data arrives. The drawbacks of the existing methods motivated us to focus on developing evolutionary clustering technique that can adapt the newly incoming

1V. Boeva, M. Angelova, E. Tsiporkova. A Split-Merge Evolutionary Clustering Algorithm, 11th International Conference on Agents and Artificial Intelligence ICAART 2019 (Prague, Czech Republic, February 19-21, 2019), accepted.

data fast without a need to re-clustering the existing clustering.

In this work, we propose evolutionary clustering techniques that are suited for applications affected by concept drift. We have developed a bipartite correlation clustering technique that can be used to adapt the existing clustering solution with new arriving data. The suggested technique is supposed to provide the flexibility to compute clusters on a new portion of data collected over a time period and to update the existing clustering solution by computed a new one. For example, some clusters will be transformed by splitting their elements among several new clusters and others will be updated by merging them with ones from arriving data. The proposed algorithm, entitled *Split-Merge Evolutionary Clustering*, is evaluated and compared to another bipartite correlation clustering technique *PivotBiCluster* on two different case studies: expertise retrieval and patient profiling in healthcare.

For future work, we are planning to extend this study with new case studies and to evaluate both algorithms on richer data sets that are being updated in different time periods.

Lilyana Boneva is a PhD student completing her dissertation at the Technical University of Sofia – branch Plovdiv. The subject of her PhD thesis is “Intelligent approaches for extraction of knowledge from Internet and other publicly available sources”. Lilyana holds a MSc degree in Computer Science from the same university. Her main research interests and activities are in the areas of artificial intelligence, bioinformatics and business intelligence. She is currently performing research dealing with data mining in the field of expertise retrieval and wearable device data analysis. Lilyana has one journal article and three papers in international conference proceedings.

ACTIVITY PROFILING OF WEARABLE DEVICE USERS

LILYANA BONEVA

Computer Systems and Technologies Department
Technical University of Sofia-branch Plovdiv
Tsanko Dyustabanov 25, 4000 Plovdiv, Bulgaria
lil2@abv.bg

We have explored the ability to profile wearable device users by applying cluster analysis techniques on biosensors’ data (passive monitoring). Device manufacturers leverage theories from psychology and behavioral economics to encourage people to wear their fitness trackers with varying success. With time, these devices have become more attractive to users. As the amount of time users spend with them grows, it becomes easier to establish a user's individual activity habits. A user activity profile can promote healthier habits as well as to serve as an early warning sign of health issues. In our study, we have used a public data set collected for a varying period of up to two years for 43 individuals. It contains data from acceleration, heart rate and skin temperature sensors. Data is sampled every second so we apply initial aggregation to form 24-hour activity vectors that we cluster for each individual with k-means. The main activity feature is acceleration, which exhibits many peaks, and the scale does not give straight answer to magnitude of the sampled activity. That is why, we have used skin temperature and heart rate, which have predetermined and known scale to normalize acceleration sensor data to a final activity vector for each 24-hour period. For each person we apply clustering analysis on her/his 24-hour activity vectors in order to capture repetitive time patterns for parts of the day or for a period. The obtained results

have shown clearly distinguishable patterns between weekends and weekdays for most of the studied individuals. However, for persons with dynamic schedules finding distinct weekday patterns has turned out to be harder. Then we cluster the individuals by k-means in order to find groups who exhibit similar activity patterns. We have applied z-standardization before performing clustering and used Dynamic Time Warping (DTW) to measure the dissimilarity between two time series. For future work, the aim is to pursue further analysis and interpretation of the clusters generated on all individuals' activity vectors.

Teodora Hristeva has a master's degree in "Information technology" from Technical University Sofia, branch Plovdiv. She has more than 15 years of experience in software projects using variety of technologies with a special impact on Java and C#. 8 years of this period she spent in Barcelona, Spain working for multinational companies and the rest in Sofia and Plovdiv, again working for big firms and also in a start-up. Currently she is a PhD student in the Computer Systems and Technologies Department with topic "Parallelizing Deep Learning algorithms using graphic accelerators". Her interests are in big data processing and analyzing.

PARALLELIZING SINCO ALGORITHM

TEODORA HRISTEVA

Computer Systems and Technologies Department
Technical University of Sofia-branch Plovdiv
Tsanko Dyustabanov 25, 4000 Plovdiv, Bulgaria
thristeva@gmail

Sparse inverse covariance selection problem is encountered in many practical applications. SINCO is an algorithm for the Sparse INverse COvariance problem. It is also an algorithm that allows parallelization.

In each step, the algorithm consists of two phases: search phase and update phase. In the search phase, the method computes potential updates for all coordinates Xij and selects the best coordinate producing the most increase in the objective function. In the update phase, using the Xij information the method updates the coordinate in the inverse covariance matrix which is a rank two update X*ij

Next step is to update the maintained estimate of using Sherman-Morrison formula for the rank-two update. This step is necessary as values of W are extensively used in the search phase of the algorithm. Each sub problem of the search phase will be a one-dimensional problem as

$$f_{ij}(\alpha) = \max \frac{n}{2} \log\det(X+\alpha(e_i e_j^T + e_j e_i^T)) - \langle S, X+\alpha(e_i e_j^T + e_j e_i^T)\rangle - \lambda|X+\alpha(e_i e_j^T + e_j e_i^T)|.$$

The algorithm is a variant of a greedy coordinate descent algorithm for a convex problem and so it enjoys known convergence guarantees and a sub linear convergence

rate. We can see that, since the updates are computed for each coordinate separately, it lends itself to distributed parallel setting. The data can be distributed among processors, assuming different distributions and methods for updating the inverse update step.

Our implementation of the parallelization is done with OpenMP and corresponds to the division of workload between different threads using the sections or parallel for directives. The first phase of the algorithm involves a simple search done on all coordinates, so we can just make the for loop parallel. In order to divide the compute load in the coordinate search phase, we need only to use #pragma parallel for for the search loop, however each thread although threads are now searching in the shared memory space, in order to save the results (maximum coordinates of each thread) we need to privatize the variables corresponding to coordinates and function value of sections corresponding to thread's territory. The important aspect that needs a specialized treatment is the update of the global best coordinate variables which will maintain the best coordinate information through all threads. To prevent race conditions we need the lines corresponding to this section in a critical statement so that no two threads execute these instructions simultaneously.

For the second phase which is updating the matrix with two rank one updates, as all the information is shared, there is no communication involved, however a barrier is needed before the second phase to ensure that all threads have synchronized information about the coordinates staged for updating and also the magnitudes of the updates. The update loop needs to be modified so that threads divide the load on updating the matrix coordinates of the inverse matrix to account for parallel environment.

The main advantage of this algorithm is the simplicity of its implementation and his parallelizing potential.

Donka Nesheva is a 3rd year PhD student at the Sofia Technical University – branch Plovdiv, Faculty of Electronics and Automation (FEA). She is currently working as an Assistant Professor at the same university. She has experience in developing Health Information Systems (HIS), Electronic Health Records (EHR) and mobile applications. Her main research interests and activities are in the area of health informatics and in particular exploring methodologies for storing and analyzing patient health data in cloud environments.

CROSS PLATFORM INTEROPERABILITY IN THE CONTEXT OF DIABETES PATIENT MONITORING SYSTEMS.

DONKA NESHEVA

Computer Systems and Technologies Department
Technical University of Sofia-branch Plovdiv
Tsanko Dyustabanov 25, 4000 Plovdiv, Bulgaria
donka.nesheva@gmail.com

Problem definition

Rapid adoption of digital technologies in healthcare has led to numerous complex distributed systems having multiple stakeholders and applications. These applications involve a multitude of platforms for managing different types of patient health data.

Last efforts in healthcare sector are focused on fully engaging patients in their self care. Data generated by smart phones and portable devices offers to the healthcare providers patients' health and lifestyle data on daily bases. Access to these data makes possible providing of personalized precision care.

However the practice of sharing patient's device generated health data has not been widely adopted by the healthcare sector. One of the reasons is the lack of successful integration of such device generated data into Electronic Health Record (EHR) or other kind of provider-accessible systems.

To make possible analytics derived from patient-generated data, it is essential to ensure data integration, which must allow working with heterogeneous data sources and platforms in interoperable ways.

Approach

My future work is focused on solving the problem in diabetes monitoring systems in particular, but could be applied to any similar use case. It addresses the interoperability issue between heterogeneous glucose monitoring systems that store patient generated

data for glucose levels, medications, diet, exercise and complains on one side and provider platforms for analysis and recommendations on the other.

I propose an approach based on development of data integration service that provides unified interface to connect and retrieve data from multiple platforms no matter the multitude of different APIs and data representations. The service encodes retrieved data according to standardized external interfaces such as Fast Healthcare Interoperability Resources (FHIR).In order to implement that service several tasks should be solved: evaluate most popular blood glucose monitoring platforms external interfaces; create API layer that unifies access to those platforms hiding their complexity; represent patient’s collected health data according to FHIR standard, which is not initially designed for that purpose.

Vasil Tsvetkov is a PhD student in his 1st year at the Technical University of Sofia – branch Plovdiv. His PhD research interests and activities are embedded systems, Internet of Things, cloud based intelligent sensor systems. Vasil holds a MSc degree in Computer Science from the same university.

USING INFRARED ARRAY DEVICES IN SMART HOME OBSERVATION AND DIAGNOSTICS

VASIL TSVETKOV

Computer Systems and Technologies Department
Technical University of Sofia – branch Plovdiv
Tsanko Dyustabanov 25, 4000 Plovdiv, Bulgaria
vaccc@abv.bg

There are many currently existing fields where infrared emission of the objects is measured, but with the increasing of the capabilities of the affordable microcontrollers and IR sensors it is possible to develop low-cost solutions for applications of infrared measurements in everyday environment. In a smart home network the IR sensor could be used as a presence device or temperature sensor.

The people detection, counting and localization are considered important and key control variables in smart sensing and modern HVAC on demand systems. Having the precision of a fraction of a degree Celsius and based on the size of the object it is possible to distinguish different objects like animals that could potentially falsely alert that there is an intruder.

However, the usage of the traditional thermal cameras that provide very high-resolution thermal images for all these applications is very expensive. It is important these systems to have reasonable cost and to be affordable for potential users in their everyday environment. In comparison of human-sensing methods for detecting presence, count, location and tracking using an inexpensive thermal infrared sensor gives promising alternative.

Having a way of obtaining a thermal image of a certain area presents a wide variety of applications in smart home, building observation and diagnostics. Very useful and attractive applications in building diagnostics are thermal loss and water leakage. The high precision measurement could detect the small temperature gradient throughout the wall, floor or ceiling and determine the source of thermal leakage. In case of water leakage IR sensor could detect the source even before the water breaks to the surface

and damages the interior.

Stefan Lishev is PhD student at the Technical university of Sofia - branch Plovdiv, Faculty of Electronics and Automation (FEA). He received B.Sc. degree in “Computer Systems and Technologies” in 2008 and a M.Sc. degree in 2010. The subject of his PhD thesis has been “System for automation of scientific experiments with remote and mobile control” but now it is changed to “System with remote and mobile access for automation of thermal fields and fluid streams measurements”. He is now an assistant professor at the Department of Control Systems. His main interests are embedded systems, artificial intelligence, digital signal processing, home automation. Stefan has four published articles in the field of his PhD thesis.

SPECIALIZED MEASURING SYSTEM FOR ANALYSING THERMAL FIELDS IN HYBRID SYSTEMS AND FLUID STREAMS

STEFAN LISHEV

Control Systems Department
Technical University of Sofia-branch Plovdiv
Tsanko Dyustabanov 25, 4000 Plovdiv, Bulgaria
stefan_lishev@tu-plovdiv.bg

Conducting research and experimentation requires controlled conditions to achieve the required precision, performance and safety. In cases where experiments are exploring slow processes, continuous measurements are required from days to months and storage of the obtained data. The subject of the dissertation is design of monitoring system with remote control. It consists of a battery powered PIC microcontroller based logging module with capability for long-term work that connects through a wireless 802.15.4 connection to a server for storage and visualization of the received data. Server database access and experiment parameter control can be done via WEB page and a mobile Android application that allows convenient and flexible management and monitoring that meets current requirements and trends in the field. The system is designed for analyzing thermal fields in phase change materials storages in hybrid systems. It is also suitable for related applications like chemical and biochemical reactors, and columns, if the temperature is in the range of -20 ° to +125 °C. Data is stored in SD-card and can be transferred via USB interface to the PC by specialized software, where the thermal fields could be visualized.

The second part of the system is for studying of the velocity field in the work chamber of the wind tunnel ULAK -1. The measurements are done using five-hole

probe in a plane located behind the subject of study. In measuring the magnitude and direction of the air flow velocity in low speed wind tunnels the probes with a spherical end portion are used. The five-hole probe is applied in this case. Central hole was located at the tip of the probe, the four side holes were symmetrically located at 45° with respect to the frontal probe axes X. For automation of the measurements and storage of data is designed special software, written in the language of LabView.

Georgi Pazhev is a PhD student in his 3rd year at the Technical University of Sofia – branch Plovdiv. His PhD research interests and activities are embedded systems, Internet of Things and smart home methods and utilities. Georgi holds a MSc degree in Computer Science from the same university.

A SURVEY OF METHODS AND TECHNOLOGIES FOR BUILDING OF SMART HOME

GEORGI PAZHEV

Computer Systems and Technologies Department
Technical University of Sofia-branch Plovdiv
Tsanko Dyustabanov 25, 4000 Plovdiv, Bulgaria
georgpajev@gmail.com

Smart home is an object for researchers from the past 50 years and the strong interest of researching continues in present days. There are several projects which are created about this topic – from smart grid infrastructure to creation of smart home projects, which are based on industrial automation and built by commercial organizations and open source hardware.

Smart homes nowadays are equipped with many devices such as smart meter, in-home displays, renewable energy sources and storage, and smart appliances such as washing machine, refrigerators, TV, oven, thermostat ,HVAC, lights, and plugs for electrical cars. A home area network is considered the backbone communication network that connects these devices. It is a two-way communication that is utilized in the demand response, advance metering infrastructure, distributed energy generations and storage. The U.S Department of Energy presents two types of home area network architectures namely: utility managed and utility and consumer managed. In utility managed architecture, the utility monitors, controls and manages smart home appliances via its private network. The smart appliances are connected to the utility control server via utility gateway and network. The user is connected to the utility via internet. He has only access to information that is provided by the utility. In the utility and consumer managed architecture the internet gateway and the utility gateway are connected in home with intermediate hub. The users and the appliances can exchange data and control commands through this common gateway/hub. Home owners can access their home appliances' control system directly through Internet gateway as well as through the utility server. Furthermore, any relevant information about the appliances can also

be delegated to a third party through the Internet.

The smart home is created because of the following main purposes – comfortability, entertainment, health care, energy management and surveillance and security. There are several technologies, which are associated with building of smart homes – IoT (Internet of Things), Cloud computing and fog computing. The protocols, which are also commonly used are CoAP (Constrained Application Protocol), MQTT (Message Queueing Telemetry Transfer), PLC (Power line Communication) and the variety sets of wireless networks like Wi-Fi, ZigBee, 6LowPan, mobile networks, etc.

There are several applications based on the technologies mentioned above. The first is an energy management system based on user centric domains, each of them contains IoT devices, which provide services to be user accessible via nomadic agent (smart phone or tablet). The next are the cloud based systems like Clome, Amazon Alexa, Google Home. The next is the IFTTT trigger action platform, which makes interaction between two IoT devices via REST API represented as channels at one recipe. At the end the protocols MQTT and CoAP are discussed.

Dimitre Kromichev is an engineer holding a Master of Science Degree in Computer Systems and Technologies. He is an assistant professor at the Department of Marketing and International Economic Relations, University of Plovdiv.

A SPEED AND ACCURACY FOCUSED APPROACH TO FPGA BASED CANNY EDGE DETECTION COMPUTATIONS

DIMITRE KROMICHEV

Department of Marketing and International Economic Relations
University of Plovdiv
24 Tzar Asen Street, Plovdiv 4000, Bulgaria
dkromichev@yahoo.com

John Canny's algorithm is the standard gauge for the contour detection segment of image processing. FPGA has been steadily increasing its market share over the last two decades.

In the literature, two main assessments stand out: 1) FPGA based Canny is too computationally complicated using exact mathematics; 2) Canny is too slow and inappropriate for real-time implementation. The objective of this study is to propose a new organization of FPGA based Canny computations guaranteeing optimal speed on the platform of total mathematical accuracy. It is accomplished by fulfilling the following tasks.

Explored is the impact of mathematical accuracy on detected contours' precision. Developed are the necessary tools and methodology. Synthetic and real life images are used to prove that approximation has a strong negative impact on precision.

Designed are five algorithms: 1) division; 2) square root; 3) inverse tangent function; 4) dynamic computation of high and low thresholds; 5) non-iterative and non-recursive hysteresis thresholding. They are explored in terms of mathematical exactness, and their performance is evaluated on a comparative basis. The following Intel (Altera) FPGA platforms are used for the tests: 130 nm – 28 nm Cyclone I – V and Stratix I – V. The algorithms are optimal in terms of speed and guarantee total mathematical accuracy.

Defined are six capital parameters of FPGA based Canny computations targeting maximum performance. Designed and explored is an entirely new approach to the organization of computations in each of the five Canny modules. Experimentally proved

is the limit of the highest clock frequency of Canny's execution in FPGA. The speed capabilities are explored on a comparative basis. For the purpose, developed is a specific methodology.

According to the plan of this study, what remains to be accomplished is: 1) Register transfer level (RTL) modelling; 2) testing detected contours' precision; 3) working FPGA based Canny contour detection system using real-life images; 4) proving the capability of the proposed organization of computations to guarantee the required positive slack for every single mathematical operation; 5) drawing relevant conclusions to the whole study.

List of Authors

List of Topics