Environmental IoT-Enabled Multimodal Mashup Service for Smart Forest Fires Monitoring

Ahmed M. Elmisery[1,*], and Mirela Sertovic[2]

Abstract

Internet of things (IoT) is a new paradigm for collecting, processing and analyzing various contents in order to detect anomalies and to monitor particular patterns in a specific environment. The collected data can be used to discover new patterns and to offer new insights. IoT-enabled data mashup is a new technology to combine various types of information from multiple sources into a single web service. Mashup services create a new horizon for different applications. Environmental monitoring is a serious tool for the state and private organizations, which are located in regions with environmental hazards and seek to gain insights to detect hazards and locate them clearly. These organizations may utilize IoT - enabled data mashup service to merge different types of datasets from different IoT sensor networks in order to leverage their data analytics performance and the accuracy of the predictions. This paper presents an IoT - enabled data mashup service, where the multimedia data is collected from the various IoT platforms, then fed into an environmental cognition service which executes different image processing techniques such as noise removal, segmentation, and feature extraction, in order to detect interesting patterns in hazardous areas. The noise present in the captured images is eliminated with the help of a noise removal and background subtraction processes. Markov based approach was utilized to segment the possible regions of interest. The viable features within each region were extracted using a multiresolution wavelet transform, then fed into a discriminative classifier to extract various patterns. Experimental results have shown an accurate detection performance and adequate processing time for the proposed approach. We also provide a data mashup scenario for an IoT-enabled environmental hazard detection service and experimentation results.

Key Words: Data mashup, Forest Fires Monitoring, Multimodal data, Internet of things.

I. INTRODUCTION

Environmental monitoring is one of the areas, which attracts public concern. The advance of cloud computing and Internet of things reshaped the manner in which the sensed information is being managed and accessed. The advances in sensor technologies have accelerated the emergence of environmental sensing service. These new services grasp the significance of new techniques in order to understand the complexities and relations in the collected sensed information. Particularly, it utilizes surveillance systems to extend the sensing range, and cloud-computing environments to analyses the big amount of data collected by various IoT sensor networks in a productive form. Various kinds of sensors are being deployed in the environment as the physical foundation for most of the environmental sensing services. However, it is highly desirable to link this sensed data with external data collected from different services in order to increase the accuracy of the predictions. In regions with environmental hazards, large number of citizens makes intensive multimedia sharing about these regions using their mobile phone during their daily activities. This massive multimedia data is expected to be generated from different sources and published on various online social services such as YouTube®, Twitter® and Foursquare®. In such a situation, it is beneficial to include such data in the decision making process of environment cognition services. Since the veracity and accuracy of the collected data are important, it is necessary for the Data Mash-up service to have the capability to pre-process the extracted data to remove noise and handle missing data issues on time before transferring it to other third parties. In this context, Data Mash-up services appear as a promising tool to accumulates, treat and manage this data in an appropriate way. The data mashup [1] is a web technology that combines information from multiple sources into a single web application for specific task or request. Mashup technology was first introduced in [2] and since then it creates a new horizon for

service providers to integrate their data to deliver highly customizable services to their customers [1, 3-5]

We believe that environmental cognition services can be enriched by extensive data collection infrastructures of IoT-enabled data mashup services especially in the domain of urban environmental monitoring. IoT mashup techniques can be used to merge datasets from external IoT networks to leverage the functionalities of environmental cognition service from different perspectives like providing more precise predictions and computation performance, improving the reliability toward citizens, minimizing the impacts of environmental hazards on affected citizens, and providing an early response in cases when the event is inevitable. Due to that, Providers of the next generation environmental cognition services keen to utilize IoT - enabled data mashup services for their systems. Effective multimedia mining is an essential requirement for the IoT - enabled data mashup services, since, the extracted patterns obviously requires the integration of different multimedia contents generated from multiple IoT networks. These multimedia contents may contain random noise, which complicates the pattern discovery process. A serious decline in accuracy occurs when the noisy data is present in the pile of contents that will be processed through the data mashup techniques. Handling this noisy data is a real challenge since it is hard to be distinguished from an abnormal data, it could prevent the environmental cognition service from fully embracing the useful data extracted from the mashup service. Managing this problem will enable the IoT - enabled data mashup services to execute different recognition methods for identifying the abnormal objects in an effective manner.

In this work, we presented an architecture supporting the evolution of IoT - enabled data mashup service business model. Our approach is based on the software agent technology, because of the fundamental features of multi-agents systems such as inter-agent communication ability; autonomy and adaptability are essential requirements of our approach. We focus on stages related to multimedia contents collection and mining and omit all aspects related to environmental monitoring, mainly because these stages are critical with regard to accuracy as they involve different entities. The proposed architecture bear in mind accuracy issues related to mashup multiple datasets from IoT networks for environmental monitoring purposes. The multimedia data is collected from the various IoT devices, and then fed into an environmental cognition service that executes different multimedia processing techniques such as noise removal, segmentation, and feature extraction, in order to detect interesting patterns in urban area. The noise present in the captured images is eliminated with the help of a noise removal and background subtraction processes.

Markov based approach was utilized to segment the possible regions of interest. The viable features within each region were extracted using a multiresolution wavelet transform, then fed into a discriminative classifier to extract various patterns. The presented approach attains accuracy and preserves the aggregates in the merged datasets in order to maximize usability and attain accurate insights. In section II describes some related work. In section III we introduced a scenario related to IoT-enabled data mashup service. In section IV introduces the proposed techniques used within the environmental cognition service. In section V describes some experiments and results based on the proposed approach. Finally, Section VI includes conclusions and future work.

II. METHODOLOGY: DATA MASHUP IN IOT-ENABLED SMART FOREST FIRES MONITORING SCENARIO

We consider the scenario where the IoT-enabled data mashup (MDMS) integrates datasets from multiple IoT networks for the environmental cognition service; figure (1) illustrates the architecture supported in this work. The proposed architecture hosts an intelligent middleware for private data mashup (DIMPM), which enables connectivity to diverse IoT devices via varied sensing technologies. In doing so, the functionalities of the proposed architecture support a cloud based infrastructure for environmental cognition services. The cloud environment promotes a service-oriented approach to big data management, providing a deep learning layer for analyzing the merged data. The architecture follows a layered approach, where the bottom layer is the Environmental IoT devices, while the highest layer is the environmental cognition service.
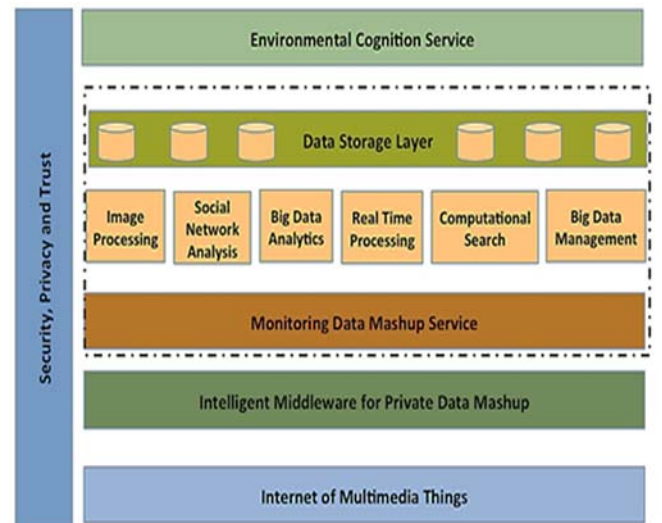

Fig. 1. IoT-enabled data mashup with Third Party Environmental Cognition Service.

The data mashup process can be summarized as follows;

- The environmental cognition service sends a query to the IoT - enabled data mashup service to gather information related to a specific region to leverage its predictions and performance.
- At the IoT - enabled data mashup:
  - The coordinator agent at search in its cache to determine the providers which could satisfy this query, then it transforms the query into an appropriate sub-queries languages suitable for each provider's database.
  - The manager agent unit sends each sub-query to the candidate IoT providers to incite them about the data mashup process.
- Based on prior agreement between the mashup provider and data providers, The providers who agree to offer purpose specific datasets to the mashup process will:
  - Forward the sub-query to its manager agent within the intelligent middleware for private data mashup.
  - The manager agent rewrites the sub-query considering the privacy preferences for its host and produces a modified sub-query for the data that can be published. This step allows the manager agent to audit all issued sub-queries and prevent ones that can extract sensitive information.
  - The resulting dataset is concealed to hide real data using the appropriate obfuscation algorithm depending of the type of multimedia data.
  - Finally, each provider submits its concealed data to the IoT - enabled data mashup service that in turn unites these results and performs further analysis on them.
- The obtained information is delivered to environmental cognition service. The environmental cognition service uses these datasets to accomplish its data analytics goals.

2.1. Proposed Anomaly Detection in the Environmental IOT-Enabled Data Mashup Service

In this section, we proposed a new service for anomaly detection with markov based segmentation approach to detect possible regions of interest, then fed the extracted features within each region into a discriminative classifier to extract various patterns. Figure 2 depicts the basic flowchart of our approach, which consists of four modules. Firstly, the noise present in the captured images is eliminated with the help of a noise removal and background subtraction processes. The second module executes Markov based approach to segment the possible regions of interest. The third module extracts the viable features within each region using a multiresolution wavelet transform. Finally, the last module, is a discriminative classifier that learns

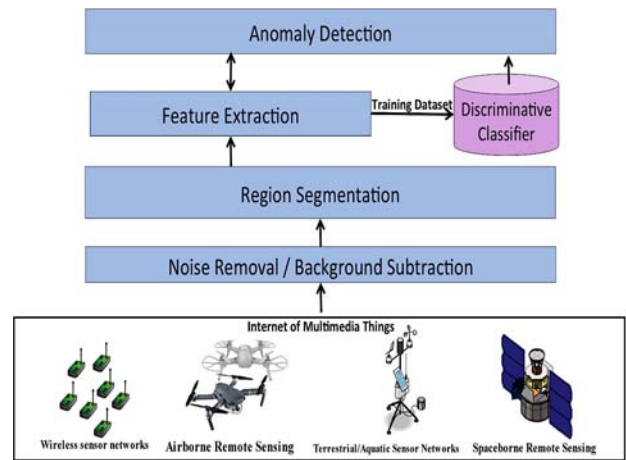effective features from each region and distinguish it into anomaly or normal region.

Fig. 2. Building Blocks of the Anomaly Detection System.

In the next sub-sections, we introduce the various steps involved in our proposed anomaly detection service. Each step utilizes an effective technique, which plays an important role in the system. The building blocks of the proposed systems are depicted in the figure (2). The proposed anomaly detection system consists of following steps:

Step 1: Noise Removal and Background Subtraction Processes

The noise present in the captured images is eliminated with the help of anisotropic diffusion combined with non-local mean and Gaussian background process [6]. The method successfully analyse the images according to each and every pixel present without eliminating the important features such as line, interpretations, and edges. More over the anisotropic diffusion process can effectively analyse blurred images. This process is applied to images using following equation:

$$\frac{\partial I}{\partial t} = div(c(x,y,t)\nabla I) = \nabla c . \nabla I + c(x,y,t)\Delta I \quad (1)$$

Where $div(c\ (x,y,t)\ \nabla I)$ represents the divergence operator of the diffusion coefficient $c(x,y,t)$ in relation with the image gradient operator ∇I. Based on (1) the anisotropic diffusion is applied to the image, if the pixel corrupted with the noise can be replaced using a non-local approach [7], where the similarity between pixels of the image is determined using the pixel intensity and is defined as follow:

$$v(i) = \ u(i) + \ n(i) \qquad (2)$$

Where $v(i)$ is defined as the current value of pixel i in given image I, $u(i)$ is defined as the "true" value of pixel i and $n(i)$ is defined as the noise mixed with the

value of pixel i. The noise exists in the image is analyzed according to the following assumption, that n (i) is an independent value extracted from Gaussian distribution with a variance $\sigma2$ and mean μ equal to 0. Based on that, the similarity between the neighboring pixels is defined depending on the weights $w(p, q_1)$ and $w(p, q_2)$. Then the non-local mean value of each pixel [7] is calculated as follows,

$$NL(V)(p) = \sum_{q\varepsilon V} w(p, q)V(q) \qquad (3)$$

V is defined as the image with noise, and the weights $w(p, q)$ satisfy that $0 \leq w(p, q) \leq 1$ and $\sum_q w(p, q) = 1$ and is defined as follow:

$$w(p, q) = \frac{1}{Z(p)} e^{\frac{-max(d^2 - 2\sigma^2(p,q))}{h}} \qquad (5)$$

σ is as defined as the standard deviation of the noise and $Z(p)$ is defined as a normalizing constant and is defined as follow:

$$Z(p) = \sum_q e^{\frac{-d(p,q)}{h}}[1,2] \qquad (6)$$

h is defined as the weight-decay control parameter. After that the neighborhood similarity value is defined as using the weighted value of the pixel and is calculated as follow:

$$d(p, q) = \|V(N^p) - V(N^q)\|_{2,F}^2 [1,2] \qquad (4)$$

F is defined as the neighborhood filter employed on the neighborhood's squared difference R_{sim} and is defined as following:

$$F = \frac{1}{R_{sim}} \sum_{i=m}^{R_{sim}} 1/(2 \neq i|1)^2 \qquad (5)$$

m is the distance between the weight and the center of the neighborhood filter. F provides higher values if the pixels near the neighborhood center, and provide lower values if the pixels near the neighborhood edge. Finally, these values are used to generate the final image.

The background subtraction was performed [8] using the Gaussian model. The background model has been constructed using the selective average method for eliminating the unwanted background pixel information as follow:

$$BM_N(x, y) = \frac{\sum_{m=1}^{N} I_m(x, y)}{N} \qquad (6)$$

Where $BM_N(x, y)$ is defined as the intensity of pixel (x, y) of the background model, $I_m(x, y)$ is defined as the intensity of pixel (x, y) of the mth frame of the captured video, and N is defined as the number of video frames utilized to construct the background model. The background model is defined using a Gaussian mixture model as follow,

$$p(x|\lambda) = \sum_{i=1}^{M} w_i g(X|\mu_i, \Sigma_i) \quad \forall i = 1, \ldots, M \qquad (7)$$

Where x is defined as continuous-valued data vector, w_i are defined as the mixture weights, and $g(X|\mu_i)$ are defined as the component of gaussian density functions [9], After that the probability value of each pixel is calculated,

$$\sum_{i=1}^{k} w_i N(\mu_t \Sigma_t, Z) \qquad (8)$$

N is the probability density function that has a mean vector μ and covariance Σ. w_i is defined as the weight of the ith Gaussian. The new pixel value Z_t is compared to each Gaussian, if the Gaussian weight is matched $\|Z - \mu_h\| < d\sigma_h$, then the Gaussian parameters are updated in accordance with:

$$w_{i,t} = (1 - a) * w_{i,t-1} + a * M_{i,t}$$
$$\mu_t = (1 - \rho) * \mu_{t-1} + \rho * Z_t$$
$$\sigma_t^2 = (1 - \rho) * \mu_{t-1} + \rho * (Z_t - \mu_t)^T * (Z_t - \mu_t)$$
$$\rho = a * N(\mu_t)$$

a is the learning rate for the Gaussian weight. Additionally, the unmatched pixel are eliminated using $w_{i,t} = (1 = a) * w_{i,t} - 1$. If none of the pixel matches the Gaussians weight, lowest weight pixel is replaced with Z_t . When the Gaussians values are stored in a corresponding index with a descending order, the initial values of this index will probably represent the background. After eliminating the background pixels and noise, the images are fed into the next step.

Step 2: Regions Segmentation

The segmentation approach is done using markov random field [10] to ensure the effective extraction of meaningful regions. It uses the local image feature value, prior probability, and marginal distribution value of the image. At the start, Markov random neighboring value must be defined from the image in terms of both first and second order neighboring values. Then the initial probability value for each feature value is set as 0 or 1. After that the mean and variance value of each pixel value is computed and labelled in the image. From the computed values, marginal distribution value is calculated according to the Bayes theorem. Finally, the probability value must be calculated and the pixels with similar values are grouped into the particular cluster or region. This process is repeated until the prior probability value reaches to a maximum value other than the defined one. The extracted regions are fed into the next step.

Step 3: Features Extraction

The multiresolution wavelet transform was employed for feature extraction. At first, the segmented regions are divided into sub-regions [10] in all the directions and then the key elements of the scale descriptors are selected. This step starts with applying a Gaussian filter on the image to detect the key elements. The maximum and minimum values of the edges are determined using the following equation $D(x, y, \sigma) = L(x, y, K_i\sigma) - L(x, y, K_j\sigma)$, where $D(x, y, \sigma)$ is the difference in the Gaussian image, $L(x, y, K\sigma)$ is the convolution value of the image $L(x, y, K\sigma) = G(x, y, k\sigma) * I(x, y)$, and $I(x, y)$ is the Gaussian blur value. Detecting key elements is accomplished using Taylor series, which is calculated as:

$$D(x) = D + \frac{\partial D^T}{\partial x}x + \frac{1}{2}x^T\frac{\partial^2 D}{\partial x^2}x \qquad (9)$$

From the detected key elements and their locations, each key elements is assigned magnitude $m(x, y)$ and orientation $\theta(x, y)$ in every direction as following:

$$m(x, y)$$
$$= \sqrt{\left(L(x+1, y) - L(x-1, y)\right)^2 + \left(L(x, y+1) - L(x, y-1)\right)^2}$$
$$\theta(x, y) = atan2\left(L(x, y+1) - L(x, y-1)\right), \left(L(x+1, y) - L(x-1, y)\right)$$

Based on the extracted key elements, different features can be calculated such as mean, standard deviation, entropy and variance. The extracted features are fed into the next step.

Step 4: Anomaly Detection

In the last step, the extracted features are used to train support vector machine classifier to detect anomalies from the captured videos, the training stage reduces misclassification error and increases the recognition rate. Each feature in the training dataset is represented as $D = \{(x_i, y_i)|x_i \in R^p, y_i \in \{-1,1\}\}$. The output value of this stage is defined as {1,-1}, in which 1 is represented as the normal feature and -1 denoted as the anomaly feature. Then the feature belongs to the class is defined by applying the hyper plane which is calculated as $w.x - b = 0$, where x is represented as the features exists in the training set, The normal hyper plane vector is w and hyper plane offset is b. The extreme learning neural networks [11] were utilized to reduce the maximum margin classification, which in turn improves the anomaly detection process. At the testing stage, the extracted features are matched with the training features to successfully detect the anomaly features. The accuracy of the proposed system was examined using the experimental results.

III. EXPERIMENTAL RESULTS

The proposed techniques were implemented in C++, we used message-passing interface (MPI) for a distributed memory implementation of anomaly detection service to mimic a distributed reliable components. We used Intel Core i7 CPU working at 2.2GHz and 8 GB RAM were used for the evaluation process. Since, there is no publicly available datasets for environmental hazards on the internet repositories. Therefore, we constructed our own datasets that utilizes the video footages of Forest fires dataset, which was provided by the National Protection and Rescue Directorate of Croatia, and other fire videos from an online social service such as YouTube. This dataset consists of 1020 fire and non-fire video clips. There are 130 forest fire video clips, 260 indoor fire video clips, 320 outdoor fire video clips and 310 non-fire video clips among the collected dataset. The resolutions of video clips were 480x360 pixels and each video clips consists of 200~300 frames. Almost 1/4 of the video clips were used for testing while the remaining were used for training. The testing set contains 40 forest fire video clips, 60 indoor fire video clips, 80 out fire video clips and 80 non-fire video clips. For negative video clips, collection of videos contains some kind of flame, such as ambulance light, Flame Effect Light, and so on. Table 1 shows that the proposed techniques achieved the real-time performance for this resolution. The most time-consuming part was related to the calculation of pixel intensity. To check the effect of the resolution of videos on the processing time, another video sequence, which had 1280x720 pixels resolution, was tested. Tests have shown that the complexity was increased by more than 2 times for 5.33 larger frame size. The Precision of this solution applied on these videos is about 94%.

Table 1: Performance of Proposed solution.

Precision	Recall	True Negative Rate	Accuracy	F-Measure	Processing Time (MS)	
					480x 360	1280 x720
0.8998	0.89	0.82	0.9543	0.9128	25.1	55.2

IV. CONCLUSION

In this work, we presented our ongoing work on building an anomaly detection in IoT - enabled data mashup service to serve environmental cognition service. A brief overview over the mashup process was presented. A Novel anomaly detection solution was also presented in detail, which achieves promising results in terms of performance.

The experiments were conducted on a real dataset and it shows that the accuracy of our solution is more than 94%. However, the ability to detect environmental hazards such as fires and reduce false positives depends mainly on the image quality. Many challenges were realized in building a data mashup service. As a result we have focused on anomaly detection within an environmental cognition service scenario. This allows us to move forward in building an integrated system while studying issues such as a privacy and security at a later stage and deferring certain issues such as virtualized schema and auditing to future research agenda.

Acknowledgement

REFERENCES

[1] T. Trojer, B. C. M. Fung, and P. C. K. Hung, *"Service-Oriented Architecture for Privacy-Preserving Data Mashup,"* presented at the Proceedings of the 2009 IEEE International Conference on Web Services, 2009.

[2] R. D. Hof. *Mix, Match, And Mutate*. BusinessWeek. Available: http://www.businessweek.com/print/magazine/content/05_30/b3944108_mz063.htm?chan=gl, 2005.

[3] A. M. Elmisery and D. Botvich, *"Agent based middleware for private data mashup in IPTV recommender services,"* in 2011 IEEE 16th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), pp. 107-111, 2011.

[4] A. M. Elmisery and D. Botvich, *"An Agent Based Middleware for Privacy Aware Recommender Systems in IPTV Networks,"* in Intelligent Decision Technologies: Proceedings of the 3rd International Conference on Intelligent Decision Technologies (IDT' 2011), J. Watada, G. Phillips-Wren, L. C. Jain, and R. J. Howlett, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 821-832, 2011

[5] A. M. Elmisery, S. Rho, and D. Botvich, *"A distributed collaborative platform for personal health profiles in patient-driven health social network,"* Int. J. Distrib. Sen. Netw., vol. 20, pp. 11-11, 2015.

[6] D. Tschumperlé and L. Brun, *"Non-local image smoothing by applying anisotropic diffusion PDE's in the space of patches,"* in Image Processing (ICIP), 2009 16th IEEE International Conference on, pp. 2957-2960, 2009.

[7] A. Narayanan and V. Shmatikov, *"Robust De-anonymization of Large Sparse Datasets,"* presented at the Proceedings of the 2008 IEEE Symposium on Security and Privacy, 2008.

[8] M. Piccardi, *"Background subtraction techniques: a review,"* in Systems, man and cybernetics, 2004 IEEE international conference on, pp. 3099-3104, 2004.

[9] A. Elgammal, D. Harwood, and L. Davis, *"Non-parametric model for background subtraction,"* Computer Vision—ECCV 2000, pp. 751-767, 2000.

[10] T. Ojala, M. Pietikainen, and T. Maenpaa, *"Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,"* IEEE Transactions on pattern analysis and machine intelligence, vol. 24, pp. 971-987, 2002.

[11] G. Feng, G.-B. Huang, Q. Lin, and R. Gay, *"Error minimized extreme learning machine with growth of hidden nodes and incremental learning,"* IEEE Transactions on Neural Networks, vol. 20, pp. 1352-1357, 2009.

Authors

Ahmed M. Elmisery is currently working as an assistant professor at the electronic engineering department of Federico Santa María Technical University (Chile). He is also a research fellow at Internet of Things and People Research Center, Malmö University (Sweden), and Adjunct Assistant Professor at Computer Science Department, Technical College (Egypt). He visited the Center for Advanced Technology in Telecommunications and Secure Systems at Monash University, Australia from May 2014 to June 2015. Before that, he was working as a Researcher in computer security at Telecommunications Software and Systems Group, Department of Computing, Mathematics and Physics, Waterford Institute of Technology, (Ireland). He received his B.S. degree in computer science from the Faculty of Computer Science, Mansoura University, Egypt (2001), M.S. degree in computer science from the Arab Academy for Science & Technology, Egypt (2007), and Ph.D. degree in computer science from Waterford Institute of Technology, Ireland (2014). He has published over 40 research papers in national and international conferences. His research interests include security, cryptography, penetration testing, malware analysis and machine learning. He is conducting research on privacy and security for future telecommunication services. His work has been grounded to develop privacy-enhanced algorithms for Internet of things.

Mirela Sertovic received the B.S. degree in Education from University of Tuzla (Bosnia and Herzegovina) in 2004. She is currently a Ph.D. candidate at University in Zagreb (Croatia). Her research interests include Social humanistic informatics, Social Networks, Blended Learning, Technology integration in teacher education, Application of technology in language learning and teaching, and Technology-assisted language learning.