FloodWatch: Flood data collection in Vietnam and LTSM-based Flood Forecasting

The impact of floods on the social and economic structure of Southeast Asia

A Thesis Prospectus

In STS 4500

Presented to

The Faculty of the

School of Engineering and Applied Science

University of Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science in Computer Engineering

By

Shuo Yan

October 27, 2022

Technical Team Members:

Ankit Gupta, Abhir Karande, Chris Santamaria, Shiva Manandhar, Shuo Yan

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

Kent Wayland, Department of Engineering and Society

Rich Nguyen, Department of Computer Science

General Research Problem: Flood Forecasting in Developing Countries

To what extent can AI change the impact of flooding on developing countries?

Floods are one of the most common and devastating natural disasters, causing widespread disruption and suffering to communities around the world. Low-income countries with underdeveloped infrastructure, such as drainage and flood protection systems, are particularly vulnerable to the destructive effects of flooding. While all countries are at risk of flooding, the majority of people at risk live in low- and middle-income countries. The frequency of flooding is directly correlated with its danger, as seen in Southeast Asia where nearly 10 million people were affected and 1,300 deaths occurred in 2011 alone (Torti, 2012). Floods can also have long-term health effects before and after their occurrence (Torti, 2012).

The extreme uncertainty of floods makes it difficult for individuals and communities to prepare in advance. Accurate prediction and warning systems are crucial in mitigating the negative impacts of flooding. Technology has advanced in this area, with the use of machine learning in flood prediction systems showing promising results. For example, Liu et al. (2018) report on the progress of China's flash flood warning systems. However, predicting floods in Vietnam and Southeast Asia may be more challenging due to the need for more data for validation due to imperfect weather monitoring. Despite the challenges, the value of accurate flood prediction in this region is immense and would have a significant positive impact on its economy.

To reduce the impacts of floods in Southeast Asia, technical research will focus on collecting flood data and developing flood prediction models through machine learning to aid in flood forecasting in Vietnam. The STS study will examine the economic and social impacts of floods in the region and the effectiveness of flood prediction systems in reducing these impacts.

Data analysis and interviews with affected communities and experts will provide insight into the needs and challenges of the region and inform the development of effective disaster risk reduction strategies.

FloodWatch: Flood data collection in Vietnam and LTSM-based Flood Forecasting

How can a flood prediction model be trained with a limited number of flood records combined with a large amount of weather data?

In collaboration with Hanoi University in Vietnam and the University of the Cote d'Azur in France, a market-oriented flood prediction application called FloodWatch will be developed using internet of things (IoT) technology and machine learning to address Vietnam's flooding problems. FloodWatch is a mobile application focused on providing flood prediction and early warning in Ho Chi Minh City and other cities in Vietnam. The application will use both research data and user-provided data to identify areas at high risk of flooding.

The development of FloodWatch will be divided into three main parts. The first part will involve training a flood prediction model using historical data, including the verification of crowdsourced flood images using real-time weather data and the identification of flood severity using neural networks such as Resnet. The second part will involve building a real-time flood prediction cloud platform and developing the FloodWatch application to provide trackable flood data for the Vietnamese user community. The third part will involve promoting the use of IoT weather condition collectors, such as long-range transmission (LoRa) devices, and arranging for the deployment of humidity and water level sensors in major cities in Vietnam. Through these steps, the FloodWatch team aims to achieve wide coverage data collection, use cloud computing for flood prediction, and provide timely flood warnings.

One of the key steps in building a flood prediction data model is the collection of weather and flood data. Weather data will be collected using weather APIs and databases. Due to the large sample size required, the team will use multithreading and asynchronous IO to accelerate the collection of geographic locations, which is expected to increase the data collection speed by 30 times (Gupta et al., 2022, 2).

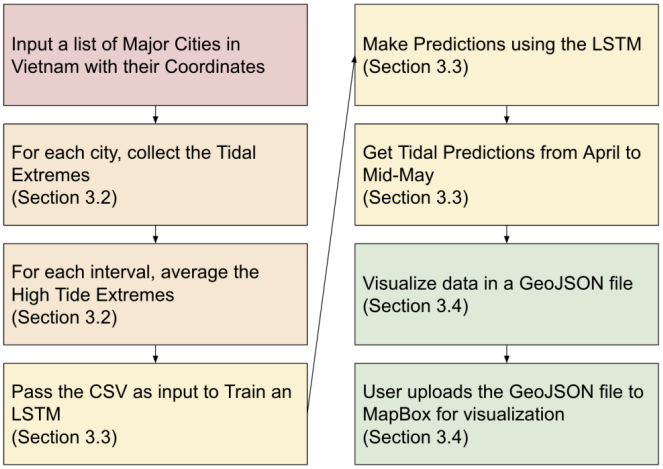

Figure 1: Methods Overview Diagram

The proposed research project will consist of several phases, as illustrated in Figure 1 (Gupta, 2022). In the initial phase, due to the limited availability of flood sample data, it will be necessary to construct a flood prediction neural network using long short-term memory (LSTM) networks. LSTM networks are particularly effective in time series-based prediction because they support the transmission and reception of signals over time. Using LSTM, a simple and effective model can be developed to predict tidal data based on time series. After training and validating the model, a key evaluation method will be to display the actual and predicted data and assess if there are any common trends.

To make the results of the research accessible and user-friendly, a user interface will be designed with a focus on fast response times. GeoJSON encoding will be used to communicate information, as it is a rich and efficient way to encode geographic data. The application will use GeoJSON encoding to assign colors to different ranges of tide values, with tide values less than 0.2 designated as "white", tide values between 0.2 and 0.4 designated as "yellow", and so on as "light orange", "orange", "light red", and "red".

All of the data and computational models developed in the project will be made available to users through a cloud service platform. The popularity of cloud platforms has made it easier to build server-based big data computing systems, and using cloud services such as AWS lambda and DynamoDB will enable automated data collection and fast prediction. As a long-term plan, building a network for collecting flood data will be necessary, which will require the cooperation of local governments and residents. This will enable the effective control of human and material losses from flooding and the reduction of the social impact of severe flooding.

While the development plan is highly implementable, there are some potential challenges. It is true that many floods can be predicted based on changes in humidity and temperature, changes in water levels, and other factors. However, due to climate change, floods are becoming more frequent and unpredictable (Chang, 2011, p. 672), which makes it necessary to train models to account for additional unpredictable factors. To obtain more informative data, flood-related information will be collected through ReliefWeb, a humanitarian information portal established by the United Nations that uses Natural Language Processing technology to analyze the severity of floods and record their locations and times.

In addition to collecting flood weather data and making predictions, the main impact of this study will be to directly assist city planners who want to be able to identify areas that are more vulnerable to flooding so that they can take appropriate action (Gupta, 2022).

The impact of floods on the social and economic structure of Southeast Asia

How do floods impact the social and economic structure of Southeast Asian countries, and what measures can be taken to mitigate these impacts?

Flooding is a frequent and destructive natural disaster in Southeast Asia, causing significant economic losses and social disruption (Tran et al., 2008). In 2011, floods in the region affected nearly 10 million people and caused over 1,300 deaths (Torti, 2012). The region is particularly vulnerable due to the concentration of the at-risk population in low-income countries with underdeveloped infrastructure systems such as drainage and flood protection (Mai & Doan, 2009). In Vietnam, for example, the lack of effective flood prevention measures has resulted in significant economic losses and disrupted the lives and livelihoods of millions of people (Tran et al., 2008). This research aims to understand the impacts of flooding on the social and economic structure of Southeast Asian countries and to identify effective measures for mitigating and managing these impacts. The findings of this study could inform the development of disaster risk reduction and adaptation strategies and support the resilience and sustainability of affected communities.

This study will use a mixed-methods approach, combining quantitative data analysis with qualitative interviews and focus groups. The quantitative data will include statistics on the frequency and economic impacts of flooding in the region, as well as survey data on the social

and psychological effects of flooding on affected individuals and communities. This data will be collected from government agencies, disaster relief organizations, and other relevant sources.

The qualitative data will include interviews and focus groups with individuals who have experienced flooding, as well as community leaders, disaster relief professionals, and government officials. The interviews will be conducted using structured and semi-structured questionnaires, and the focus groups will be facilitated by trained moderators. The data will be analyzed using both qualitative and quantitative methods, including thematic analysis, regression analysis, and network analysis. The findings will be triangulated and validated through multiple sources and methods to ensure the reliability and validity of the results.

The data will be analyzed to identify trends and patterns in the social and economic impacts of flooding, as well as potential interventions and policies for mitigating these impacts. The study will also explore the role of insurance and technology in reducing the impacts of flooding in the region, and will examine the challenges and opportunities for implementing these measures.

The expected outcomes of this study include a better understanding of the social and economic impacts of flooding in Southeast Asia and the needs of affected communities. The study may also identify potential interventions and policies for mitigating the impacts of flooding, such as early warning systems (Ren et al., 2018), improved infrastructure, and community-based disaster risk reduction programs. For example, in China, the development of flash flood warning systems using machine learning has been shown to be effective in reducing the impacts of flooding (Liu et al., 2018). In Vietnam, the implementation of community-based disaster risk reduction programs has been shown to increase the resilience of affected communities and reduce the social and economic impacts of flooding (Kappas & Vogt, 2014).

In addition to these interventions, the study may also explore the potential role of insurance in mitigating the impacts of flooding. In many Southeast Asian countries, the majority of the population is not covered by insurance, and the lack of insurance coverage increases their vulnerability to the economic impacts of flooding (Mai & Doan, 2009). However, insurance can provide a financial safety net for individuals and communities affected by flooding, allowing them to recover faster and reduce the long-term impacts on their livelihoods and well-being (Islam & Sultan, 2017). The study may examine the feasibility and potential benefits of implementing insurance schemes for natural disasters in the region, including the challenges and barriers to adoption.

Furthermore, the study may also explore the role of technology and innovation in reducing the impacts of flooding in Southeast Asia. With the increasing availability of mobile devices and network facilities, there is an opportunity to develop smart cities and connected infrastructure that can improve flood warning and response systems (Vu, 2018). The study may examine the potential of using machine learning and other advanced technologies to improve flood prediction and warning, as well as the challenges and opportunities for implementing these technologies in the region.

Overall, this research project aims to provide a comprehensive understanding of the impacts of flooding on the social and economic structure of Southeast Asian countries and to identify effective measures for mitigating and managing these impacts. The findings of this study may inform the development of disaster risk reduction and adaptation strategies, support the resilience and sustainability of affected communities, and contribute to the broader literature on the impacts of natural disasters on human development.

Conclusion

Life is fragile, the property is valuable, and the important things in life can easily be lost to unpredictable disasters. The proposed research project aims to develop a flood prediction system, FloodWatch, using machine learning algorithms in Vietnam. The technical aspect of the project focuses on the implementation of the prediction system, including the use of long short-term memory neural networks and cloud-based data collection and analysis. The STS aspect of the project examines the economic and social impacts of floods in Southeast Asia, and the potential benefits of effective flood prediction in reducing these impacts. The proposed research will provide valuable insights into the challenges and opportunities for mitigating the impacts of flooding in Vietnam and the broader Southeast Asian region. By highlighting the potential benefits of FloodWatch, the project aims to contribute to the development of effective disaster risk reduction and adaptation strategies in the region. Overall, the research aims to support the resilience and sustainability of communities in Vietnam and Southeast Asia.

References

Chang, C.-H. (2010). Preparedness and storm hazards in a global warming world: Lessons from Southeast Asia. Natural Hazards, 56(3), 667–679. https://doi.org/10.1007/s11069-010-9581-y

Gupta, A., Kim, A., Karande, A., Yan, S., Manandhar, S., & Nguyen, R. (2022, 7 30). Validating Crowdsourced Flood Images using Machine Learning and Real-time Weather Data. 2022 IEEE International Conference on Big Data Science and Engineering.

Gupta, A., & Nguyen, N. R. (2022). Predicting and Visualizing Tidal Trends in Vietnam using an Automated Geographical Data Pipeline.

Islam, M. S., & Sultan, M. (2017). Impact of insurance coverage on community vulnerability to natural disasters: A case study in Bangladesh. Natural Hazards, 87(3), 1443-1466.

Kappas, M., & Vogt, J. (2014). Community-based disaster risk reduction in Vietnam: A case study of the Mekong Delta. International Journal of Disaster Risk Reduction, 10, 66-74.

Liu, C., Guo, L., Ye, L., Zhang, S., Zhao, Y., & Song, T. (2018). A review of advances in China's flash flood early-warning system. Natural Hazards, 92(2), 619–634. https://doi.org/10.1007/s11069-018-3173-7

Mai Van, C., van Gelder, P., & Vrijling, J. (2009). Risk based design of coastal flood defenses - a Vietnam case. Reliability, Risk, and Safety. https://doi.org/10.1201/9780203859759.ch156

Ren, H.-L., Wu, Y., Bao, Q., Ma, J., Liu, C., Wan, J., Li, Q., Wu, X., Liu, Y., Tian, B., Fu, J.-X., & Sun, J. (2019). The china multi-model Ensemble prediction system and its application to flood-season prediction in 2018. Journal of Meteorological Research, 33(3), 540–552. https://doi.org/10.1007/s13351-019-8154-6

Torti, J. (2012). Floods in Southeast Asia: A health priority. Journal of Global Health, 2(2). https://doi.org/10.7189/jogh.02.020304

Tran, P., Marincioni, F., Shaw, R., Sarti, M., & An, L. V. (n.d.). Flood Risk Management in central viet nam: Challenges and potentials. Keio University. Retrieved December 7, 2022, from https://keio.pure.elsevier.com/en/publications/flood-risk-management-in-central-viet-nam-challenges-and-potentia

Vu, K., & Hartley, K. (2018). Promoting smart cities in developing countries: Policy Insights from Vietnam. Telecommunications Policy, 42(10), 845–859. https://doi.org/10.1016/j.telpol.2017.10.005