A Novel Cloud-Based Crowd Sensing Approach to Context-Aware Music Mood-Mapping for Drivers

Arun Sai Krishnan[1,2], Xiping Hu[2], Jun-qi Deng[3], Renfei Wang[4], Min Liang[5], Chunsheng Zhu[2], Victor C.M. Leung[2], Yu-Kwong Kwok[3]

[1] Department of Computer Science and Engineering, National Institute of Technology, Tiruchirappalli, India

[2] Dept. Electrical and Computer Engineering, The University of British Columbia, Vancouver, Canada

[3] Electrical and Electronic Engineering Department, The University of Hong Kong, Hong Kong

[4] IBM, Ottawa, Canada

[5] IBM, Guangzhou, China

Email: arunsaikrish@gmail.com, {xipingh, cszhu, vleung }@ece.ubc.ca, {jqdeng, ykwok}@eee.hku.hk, renfeiw@ca.ibm.com

***Abstract*— Millions of people are severely injured or killed in road accidents every year and most of these accidents are caused by human error. Fatigue and negative emotions such as anger adversely affect driver performance, thereby increasing the risk involved in driving. Research has shown that listening to the right kind of music in these situations can ameliorate driver performance and improve road safety. Context-aware music delivery systems succeed in delivering suitable music according to the situation through the process of music mood-mapping which identifies the mood of a song. Additionally, we can leverage the power of the cloud to enable crowd sensing of the mood-mapping of various songs and enhance the effectiveness of situation-aware music delivery for drivers. The cloud can be used to aggregate the crowd sensed music mood-mapping data and improve the effectiveness of music delivery by providing accurate mood-mappings from the aggregated data. Currently, context-aware music delivery systems consider only features from the song for music mood-mapping. In this paper, we propose a novel approach to music mood-mapping for drivers which also incorporates the social context of a driver including age, gender and cultural background to enhance the effectiveness of music delivery in context-aware music recommendation systems for drivers.**

Keywords*— music mood-mapping; context-aware; cloud; crowd sensing; safe driving

I. INTRODUCTION

Statistics published by the World Health Organization (WHO) in the report on road traffic injury prevention [1], reveal that the total number of deaths due to road accidents is extremely high at 1.2 million per year. Thus, promotion of safer driving and road safety are key to reducing the numbers and making our roads safer. A study from Europe [2] shows that reduced vigilance of drivers arising from fatigue, stress and negative emotions such as anger result in approximately 10-20% of all road accidents. Music, with its power to ameliorate the mood of the listener, can be used as a powerful tool to enhance the mood of a driver in case of fatigue or stressful situations. Previous research [3] has shown that listening to suitable music while driving does not distract the driver, but rather enhances driving performance. Thus, context-aware music delivery systems for drivers can promote safer driving as effective use of suitable music will lead to enhanced driving performance and safer roads.

Traditionally, context-aware music delivery systems for drivers utilize the process of music mood-mapping to identify the mood of a song. A music label is obtained for every song in the music library of a driver's smartphone through the process of music mood-mapping. The current mood of a driver can be analyzed through data from various sensors attached to a smartphone such as the camera and heartbeat monitor. Based on the current situation and mood of the driver, the context-aware music delivery system chooses the song with the most suitable music label for the current situation of the driver.

Efficiency and effectiveness are two critical factors related to context-aware music delivery system for drivers. The system should recommend suitable music to the driver quickly in order to ensure that the driver listens to the appropriate music before the context changes. Hence, the system needs to be efficient in order to ensure timely delivery of suitable music. Effectiveness of music delivery deals with delivery of suitable music for a given situation. For instance, soothing music should be played when the driver is angry and energetic music should be played when the driver is experiencing fatigue.

In this paper, we present our novel cloud-based crowd sensing approach to improve the effectiveness of music mood-mapping. Traditional music mood-mapping methods consider only music features extracted from the song while identifying the mood of a song [4] [5]. However, different people might experience different moods for the same song. The social context of a driver, including data regarding the driver's age, gender and cultural background affects the driver's perception of the mood of a song [6] [7] [8]. Hence, our solution also takes into account the social context of the driver and meta-data from the song while performing the music mood-mapping process in order to improve the effectiveness of music delivery and provide accurately customized mood-mappings based on the social context of the drivers. Different songs from the same artist tend to have a lot of similarities which may not be captured by the numerical values of the extracted music features. Data regarding the artist of the song can be obtained from the meta-data of a music file. We use a sequence-matching based similarity detection process to identify songs which have the same artist. We use this similarity detection technique while clustering the mood-mappings of all songs in the cloud, which is explained in detail in the next section.

II. DESIGN

Our framework for cloud-based crowd sensing of music mood-mapping for drivers mainly consists of two tiers named the mobile tier and the cloud tier. A detailed illustration of the architecture of our framework is presented in Fig. 1.

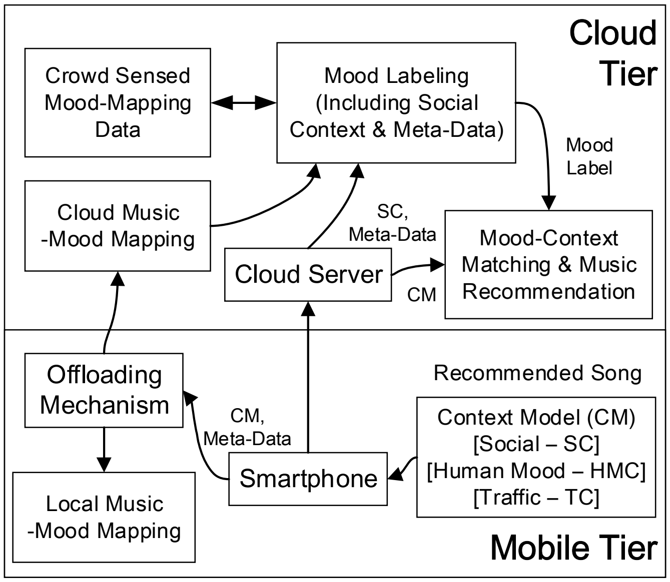

Fig. 1. Architecture Diagram

The situation-aware music delivery system resides in a smartphone which is a component of the mobile tier. The smartphone interfaces with the various sensors from which the context-aware music delivery system derives contextual data in order to identify the context of the driver. Using this contextual data, the context model (CM) consisting of social context (SC), human mood context (HMC) and traffic context (TC) is formed. Music mood-mapping is a computationally intensive process and thus requires considerable computing time and power. Cloud systems are generally much more powerful than smartphones and can thus perform music mood-mapping comparatively quicker. However, transferring an entire music file to the cloud for mood-mapping involves considerable network overheads and data transfer time. Hence, in some situations performing mood-mapping locally might be quicker than offloading it to the cloud. The situation-aware music delivery system consists of an offloading mechanism which decides whether to perform music mood-mapping of a song locally or offload the computation to the cloud.

Utilizing the cloud based system in our system has its advantages. Firstly, it enables the crowd sensing of music mood-mapping which helps in accurately determining the mood of a song, as explained in the following sections of this paper. Secondly, the enormous computing power of the cloud can be leveraged to reduce mood-mapping time and improve efficiency of the situation-aware music delivery system in many cases. The cloud server is the first component of the cloud tier. It receives data from the smartphone and choreographs the communications between the other components of the cloud tier. The cloud music mood-mapping component is used to receive the offloaded song and perform mood-mapping if the offloading mechanism decides to offload mood-mapping to the cloud. In case mood-mapping is not offloaded, this component just receives the mood-mapping computed locally on the smartphone and passes it on to the mood labeling component. The role of the mood labeling component is to accurately identify the mood of a song and label it based on the crowd sensed mood-mapping data of various songs accumulated in the cloud. In the currently existing situation-aware music delivery systems, this process of music mood labeling is done solely using musical features extracted from the songs. As explained in the next subsection, our approach to music mood labeling incorporates the social context of the driver including age, gender and cultural background in the process in order to improve effectiveness of music delivery and provide accurately customized mood-mappings based on the driver's social context. The final component in the cloud tier is responsible to match the current context model to an appropriate music mood label in order to recommend the right song for the given situation.

A. *Mood Mapping*

In our approach, we adopt the skeletal version of the mood-mapping process from SAfeDJ [5] where six musical features of a song namely zero-crossing rate, unit power, low energy rate, tempo, spectral centroid and tonal type are extracted from a song and used as an indicator of their mood. This mix of temporal and spectral music features gives a good indication of the mood of a song.

B. *Mood Labeling*

Mood labeling is the process which gives a mood label to a song and this mood label is what identifies a song with a mood. Mood labels can be obtained from feelings associated with a song. A song can be lively, energetic, sad, groovy, noisy or peaceful. It can also be a combination of all of these. As suggested in the SAfeDJ [5] system, the mood label can be a point in the six dimensional space of these six moods. There is an underlying relation between these six mood dimensions and six musical dimensions as established in SAfeDJ. Initially, 21 representative songs are chosen to represent different moods and based on a survey the six dimensional mood label for each of these representative songs is determined as in SAfeDJ. When a new song is then mood-mapped, a clustering algorithm is applied as adopted in SAfeDJ. The new song is added to the cluster of the point nearest to it in the six dimensional mood space and is assigned the label of the centroid of that cluster if the distance is lesser than the defined distance threshold. If the distance to the nearest point is greater than the distance threshold, then a new cluster is created and the mood label is the assigned the co-ordinates of that point. In the traditional system, the distance measured is the Euclidean distance and hence gives equal weightage to all the six mood dimensions. However, based on different social contexts of people such as difference in age, gender and culture, their respective perception of the different mood dimensions in a song will be different and hence assuming equal weightage to the different mood dimensions while calculating the distance will give inaccurate results. Thus, to solve this problem we use a custom weighted Euclidean distance to measure distance between points during the clustering process. We've also incorporated a similarity detection system based on Gestalt pattern matching that identifies songs from the same artist using meta-data from the song. The distance in the six dimensional mood space between songs with the same artist is then reduced accordingly in order to account for the inherent similarity between them as they are from the same artist. This process helps in taking into consideration inherent similarity between songs, which is not captured by numeric values of the musical features.

In our approach, we conduct a survey in which people are asked to pick the different moods they feel while listening to each song in a set of songs. The social contexts such as age, gender and cultural background of the participants in the survey are recorded as well. The data obtained from this survey is used as a baseline and a true indicator of the mood label of a song for different social contexts. Each of these songs is then mood-mapped through the mood-mapping process. Then, data from all the participants are grouped according to social context. Each of these groups has a unique social context. Then, the voted music labels for the set of all songs in each social context is calculated. Then, the average difference between mood-mapping and voted value of each mood dimension in all of the songs for a given social context is calculated. Let them be d_L, d_E, d_S, d_G, d_N and d_P where L, E, S, G, N and P stand for lively, energetic, sad, groovy, noisy and peaceful respectively. The total difference in all the moods is obtained by summing up the difference in each mood dimension. The contribution of one mood dimension to the total difference is measured as a ratio as shown below.

$$r_L = \frac{d_L}{d_L + d_E + d_S + d_N + d_G + d_P} \qquad (1)$$

The weight of each mood in the calculation of the distance is then obtained using the inverse of the above ratios. For instance, if $r_L > r_E > r_S > r_N > r_G > r_P$, then $w_L = r_P$, $w_E = r_G$, $w_S = r_N$ and so on. Finally, the distance between any two points in the six dimensional mood space is calculated using our custom weighted Euclidean distance as shown below.

$$Dist = \sqrt{6\sum_{i=L}^{P} w_i (x_i - y_i)^2} \qquad (2)$$

This same process is then repeated across all the different social contexts in order to obtain the optimal weight distribution amongst the mood dimensions for each social context, thereby resulting in different and customized mood weight distribution for each social context. Thus, we incorporate the social context of a driver as well into the music mood-mapping process through the use of our non-equal weight distribution amongst moods and custom weighted Euclidean distance measure. This process improves the effectiveness of music mood-mapping and also enables social context aware music mood-mapping which provides customized mood-mapping based on the social context of the driver such as age, gender and cultural background.

III. Initial Results and Implementation

Based on our implemented platform SAfeDJ [5] [6] 10], we conducted a survey among 80 people in which we picked 6 songs and asked them to pick what moods out of the six mood dimensions they could feel while listening to each song. Then, we applied music mood mapping to these songs and calculated the music mood-mapping (MM) of these songs. Also, We aggregate the voting data in groups based on the social context of the participants. Then, the voted music label (VM) is calculated from the aggregated data. This process uses the mean of all the individual voting data from the participants of a particular social context in order to calculate the VM. For each

Name Of Song	Lively (L)	Energetic (E)	Sad (S)	Groovy (G)	Noisy (N)	Peaceful (P)	Distance (Original Method)	Distance (Our Approach)
	$\lvert MM-VM \rvert = d_L$	$\lvert MM-VM \rvert = d_E$	$\lvert MM-VM \rvert = d_S$	$\lvert MM-VM \rvert = d_G$	$\lvert MM-VM \rvert = d_N$	$\lvert MM-VM \rvert = d_P$		
Never Wake Up	0.035	0.162	0.0	0.274	0.106	0.0	0.337	0.331
Waltz For Debby	0.061	0.0	0.696	0.078	0.030	0.056	0.706	0.465
Huohuo	0.488	0.048	0.304	0.236	0.030	0.254	0.674	0.554
Fuzz Universe	0.177	0.383	0.0	0.152	0.063	0.0	0.453	0.460
And I Love Her	0.094	0.078	0.304	0.114	0.121	0.200	0.419	0.370
Nizaiganma	0.257	0.030	0.391	0.169	0.030	0.054	0.502	0.364

Original Method Average Distance = 0.515

Our New Method Average Distance = 0.424

Improve = 17.67%

Fig. 2. Experimental Results and Comparison

of the social context based groupings, we calculate the optimal mood weight distribution as described in the previous section.

The results obtained by utilizing our novel approach are illustrated in Fig. 2. The data in this figure is presented from the perspective of the social context of an Asian male who is aged between 20 and 30 years. We can compute and showcase the results of other social contexts as well in a similar way. We first calculate the distance between the VM and MM using the method to calculate normal Euclidean distance which is used in existing systems for context-aware music delivery for drivers. Then, we calculate the distance between the MM and VM using our custom weighted Euclidean distance measure, with appropriately calculated weights for the social context of an Asian male who is aged between 20 and 30 years.

It is clearly seen from the values in the table that distance between the VM and MM obtained through our novel approach is considerably lesser than the distance obtained by the traditional approach. This shows that our novel approach is successful in improving effectiveness and accuracy of the music mood-mapping process by incorporating the social context of a driver in the music mood-mapping process. Also, the average distance between the VM and MM in the traditional method was 0.515, but using our approach, it has been reduced to 0.424. Hence, our novel solution has improved the accuracy of music mood-mapping by 17.67%. This is a really good improvement over the existing method and the initial results we've obtained look promising.

In addition, we used native Java and the Android platform for implementing the mobile tier of our solution. The cloud tier was developed using Python-Flask and Java. Analysis of survey and experimental data was done using Python scripts.

IV. Conclusions and Future Work

Road safety is a major cause of concern due to the unacceptably high number of fatalities and serious injuries caused by road accidents every year. Fatigue, stress and negative emotions such as anger experienced by a driver are shown to be major causes of road accidents. Situation-aware music delivery systems for drivers leverage the powerful ability of music to ameliorate the effects of fatigue, stress and negative emotions amongst drivers. Delivering the right song for a given situation is a very crucial part of these systems which use music mood-mapping to identify the mood of a song. Suitable and appropriate music should be delivered according to the situation in order to improve driver performance. Thus, arises the need for improving the effectiveness of the music mood-mapping process for drivers. Traditional context-aware music delivery systems for drivers utilize just the musical features of a song and a Euclidean distance measure while clustering songs based on moods in the music mood-mapping process. However, different people might have different mood perceptions for the same song based on their social context which includes their age, gender and cultural background. The social context of a driver also plays a crucial role in the mood perception of a song by the driver and hence the traditional method falls short in this aspect as it doesn't take into account the social context of a driver in the music mood-mapping process. In this paper, we've described our novel cloud-based crowd sensing approach to improve the effectiveness of music mood-mapping for drivers. Our approach takes into consideration the social context of a driver in the music mood-mapping process and utilizes a custom weighted Euclidean distance measure for the music mood-mapping process. We also consider the meta-data from the music in the mood-mapping process in order to account for similarity between songs, which might not be captured by the numeric values of the musical features. Hence, our comprehensive solution improves effectiveness and accuracy of music mood-mapping considerably.

In the future, we plan to improve the accuracy of our novel approach by conducting a survey with a larger song set that represents more moods and with more participants from varied social contexts. We also plan to include more factors from the social context and meta-data in order to further enhance the accuracy and effectiveness of our novel approach for music mood-mapping for drivers.

V. Acknowledgements

This work is supported in part by Mitacs Globalink, the Canadian Natural Sciences and Engineering Research Council through the NSERC DIVA Strategic Research Network, TELUS and other industry partners.

References

[1] World Health Organization (2014). *World report on road traffic injury prevention*. [Online]. Available: http://www.who.int/violence_injury_prevention/publications/road_traffic/world_report/en

[2] European Commission (2015). *A study by Transport - European Commission.* [Online]. Available: http://ec.europa.eu/transport/road_safety/specialist/knowledge/fatique/index_en.htm

[3] M. Zwaag, C. Dijksterhuis, D. Waard, B. Mulder, J. Westerink, and K. Brookhuis, "The influence of music on mood and performance while driving," *Ergonomics*, vol. 55, no. 1, 2012, pp.12-22.

[4] P. Helmholz, S. Vetter, and S. Bissantz, "AmbiTune: Bringing Context-Awareness to Music Playlists while Driving," *Advancing the Impact of Design Science: Moving from Theory to Practice.* 8463, Springer International Publishing, 2014, pp. 393-397.

[5] X. Hu, J. Deng, J. Zhao, W. Hu, E. C.-H. Ngai, R. Wang, J. Shen, M. Liang, X. Li, V. C. M. Leung, and Y. Kwok, "SAfeDJ: A Crowd-Cloud Co-design Approach to Situation-aware Music Delivery for Drivers," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 12, no. 1s, 2015.

[6] X. Hu, J. Deng, W. Hu, G. Fotopoulos, E.C-H. Ngai, Z. She ng, M. Liang, X. Li, V.C.M. Leung, and S. Fels, "SAfeD J Community: Situation-Aware In-Car Music Delivery for Safe Driving," *In Proc. ACM MobiCom*, 2014, pp. 363-366.

[7] G. Soley, and E. Hannon, "Infants prefer the musical meter of their own culture: A cross-cultural comparison," *Developmental Psychology*, vol.46, no.1, Jan 2010, pp. 286-292.

[8] W. Thompson, and L. Balkwill, "Cross-Cultural Similarities and Differences, in Handbook of Music and Emotion," *Oxford University Press*, 2010, ch. 27, pp. 755–788.

[9] T. Teo, D. Hargreaves, and J. Lee, "Musical preference, identification, and familiarity: A multicultural comparison of secondary students from Singapore and the United Kingdom," *Journal of Research in Music Education*, 56, 2008, pp. 18-32.

[10] Xiping Hu. (2015). *Context-aware mobile crowdsensing in mobile social networks*. [Online] Available: www.mobilesoa.appspot.com