Entrain: Encouraging Social Interaction in Collective Music Making

Hugo Scurto
Wanyu Liu
Benjamin Matuszewski
Frédéric Bevilacqua
STMS IRCAM–CNRS–Sorbonne Uni.

Jean-Louis Frechin
Uros Petrevski
Nodesign.net

Norbert Schnell
Furtwangen University

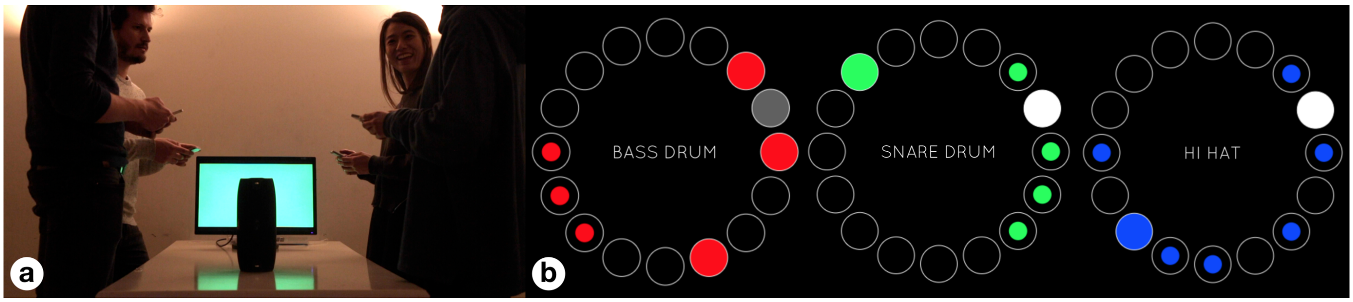

Figure 1: (a) Collective music making using *Entrain* in *Coloop* and (b) UIs for 3 users with their respective digital instrument.

ABSTRACT

Entrain is an adaptive agent designed to stimulate social interaction among users in collective music making. It registers individual user behavior and provides sonic feedback to encourage users to look up from their mobiles and interact with each other. We demonstrate a use case of *Entrain* in *Coloop*, a distributed and collective musical instrument.

CCS CONCEPTS

• **Human-centered computing** → *Collaborative and social computing devices.*

KEYWORDS

Social, Mobile, Adaptive, Audio/Music, Bayesian information gain

ACM Reference Format:
Hugo Scurto, Wanyu Liu, Benjamin Matuszewski, Frédéric Bevilacqua, Jean-Louis Frechin, Uros Petrevski, and Norbert Schnell. 2019. Entrain: Encouraging Social Interaction in Collective Music Making. In *Proceedings of SIGGRAPH '19 Studio.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3306306.3328004

1 INTRODUCTION

Collective music-making systems have gained ongoing popularity in recent years, notably through advances in mobile and web technology. They leverage the ubiquity and multi-modality of these devices for both musicians and non-musicians to create, compose and share music collaboratively. Furthermore, mobile musical systems create novel co-located audiovisual experiences, which offer the possibility of accessible and engaging group activities.

Collective Loops[1] and *Coloop*[2] are examples of these collective music-making systems in public spaces. They allow individual user to generate rhythmic sound by interacting with his/her mobile phone and synthesize them in real time. However, two challenges emerged after an initial installation of these systems. First, even if they support simultaneous multi-user interaction toward a shared outcome, most participants remained individual. This has been found in many previous studies concerning public displays that many people have troubles engaging in a social context if they do not know each other [Yoon et al. 2004]. Second, the lack of formal musical education hinders non-musician participants from fully exploring the potential music creation space. For instance, in *Coloop*, many participants focused on filling out all the circles (Fig. 1), rather than listening to each other and finding the rhythmically sound combination of these circles.

We introduce *Entrain*, an adaptive agent designed to tackle the two above-mentioned challenges. We are inspired by the notion of *musical entrainment*, which refers to the human phenomenon of rhythmic synchronization that may occur when listening or playing music [Varni et al. 2011]. While users are engaged in collaborative acts of music making, *Entrain* registers user actions, encourages and guides music creation at both individual and collective level. Individually, it rewards the user based on his/her individual musical behavior; collectively, it adapts its rewarding criterion based on all users' behavior. This is done by leveraging the framework of

SIGGRAPH '19 Studio, July 28 - August 01, 2019, Los Angeles, CA, USA

ACM ISBN 978-1-4503-6316-7/19/07.
https://doi.org/10.1145/3306306.3328004

[1] http://cosima.ircam.fr/2015/11/25/collective-loops-forum-workshops
[2] https://www.nodesign.net/en/portfolio/coloop

Bayesian Information Gain [Liu et al. 2017] and introducing special audiovisual effect extracted from music features. We introduce the detailed implementation of *Entrain* in *Coloop* in the next section.

2 ENTRAIN IMPLEMENTATION

Entrain builds on *Coloop* sequencer implementation, which can connect up to eight mobile users together and provide them with 16-bit sequences corresponding to specific percussive elements (see Fig. 1 (b)). It leverages our previous framework for collective mobile web interaction [Schnell et al. 2017] to give each sequence a common reading head and synchronize the eight layers of sequences. Importantly, each sequence is looped indefinitely.

2.1 User State Features

To describe users state, we designed two musical features relevant to the connected sequencer use case (see Fig. 2). The first feature is *user activity* and is based on the number of changes made by the user in the sequence in each loop. A completely passive user would have null activity, while an overactive user would have maximum activity. The second feature is *rhythmic periodicity* and is based on the autocorrelation[3] of the user's sequence. A fully-filled sequence would have minimum periodicity, while a large-spaced sequence would have maximum periodicity. Both features are computed at the end of each loop, which supports dynamic accounting of a user's state as he/she interacts with the system. Note that our approach is general; one may use other features—and more than two—to describe a user's state in other collective use cases.

2.2 Agent Adaptation

We use the framework of Bayesian Information Gain (BIG) [Liu et al. 2017] to support machine adaptation (Fig. 2). BIG allows to adapt a probabilistic model of user optimal behavior $p(\Theta = \theta)$ by actively sending some feedback $X = x$ to users and sensing users' subsequent input $Y = y$.

For *Entrain*, we modified the framework to the scenario of connected sequencer. In particular, we used it to model the user state feature space. $p(\Theta = \theta)$ corresponds to the probability that a given *user activity* and *rhythmic periodicity* may correspond to the user's optimal behavior. It is possible to define prior knowledge if we want the agent to encourage users to adopt a specific state (e.g., medium activity with high rhythmic periodicity). Thus, the user's goal becomes to find the optimal state of the model.

To find the optimal state of the model, users have to explore different levels of activity and rhythmic periodicity. As previously stated, this may be a hard task if users are not trained musicians. *Entrain* thus sets a temporary target state in the user state space, corresponding to X in the BIG framework. In the case where a user gets into the target state, he/she is rewarded by the agent at a multimodal level. The sound produced by his/her instrument will be augmented with an audio effect, while the color corresponding to his/her instrument will be displayed on a shared screen (see Fig.1 (a)). The multimodal reward thus acts as a social incentive: it encourages the rewarded user to continue performing at the same level, while motivating other users to play along with him/her to benefit from the same reward.

[3]https://en.wikipedia.org/wiki/Autocorrelation

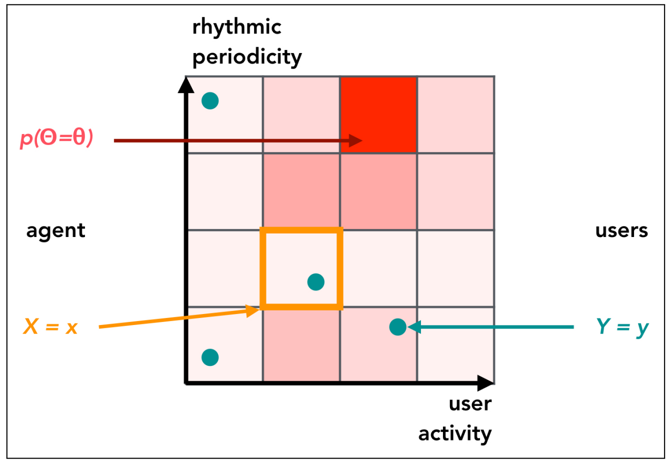

Figure 2: *Entrain* implementation. The two axes corresponds to user state features. The 4 dots represent the state of 4 different users. $P(\Theta)$, X, and Y correspond to the parameters in the BIG framework for adaptation.

After some time of interaction, the agent may update its knowledge on users' optimal goal. The update is done by measuring users' move in the feature space, which corresponds to Y in the BIG framework. The agent updates its probabilistic model of user optimal behavior and generates a new temporary target state X' that will be active for the next eight loops. In the case where no user gets into the target state after eight loops, the agent will take two strong actions toward users. First, it stops interaction during four loops to reward the user that is closest to the temporary target state. This allows users to switch their attention from their individual interfaces to the collective outcome. Second, it will slightly adjust user states toward the temporary target state. This may help users finding optimal state and obtaining reward.

3 CONCLUSION

We have presented an adaptive agent for encouraging social interaction in collective music making. Future work includes field studies with users in public spaces, as well as others applications leveraging mobile interaction and sonic feedback.

REFERENCES

W. Liu, R. L. d'Oliveira, M. Beaudouin-Lafon, and O. Rioul. Bignav: Bayesian information gain for guiding multiscale navigation. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 5869–5880. ACM, 2017.

N. Schnell, B. Matuszewski, J.-P. Lambert, S. Robaszkiewicz, O. Mubarak, D. Cunin, S. Bianchini, X. Boissarie, and G. Cieslik. Collective loops: Multimodal interactions through co-located mobile devices and synchronized audiovisual rendering based on web standards. In *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '17, pages 217–224, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4676-4. doi: 10.1145/3024969.3024972. URL http://doi.acm.org/10.1145/3024969.3024972.

G. Varni, M. Mancini, G. Volpe, and A. Camurri. A system for mobile active music listening based on social interaction and embodiment. *Mobile Networks and Applications*, 16(3):375–384, 2011.

J. Yoon, J. Oishi, J. Nawyn, K. Kobayashi, and N. Gupta. Fishpong: encouraging human-to-human interaction in informal social environments. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 374–377. ACM, 2004.