suit of this goal: how much should the system lead the musicians to help them stay in time without making the performance artificial? Predicting musical timing with sufficient accuracy will open up interesting avenues for network music research, especially when we consider parallel research into predicting other information such as intensity and even pitch information, but whether any musician would truly want to let a machine impersonate them expressively remains to be seen, which is why we propose that a 'minimally-invasive' conductor-like approach to regulating tempo would be more appropriate than complete audio prediction.

5.1 The Bayesian Network

It would be straightforward to extend our model by implementing prediction of timing from other forms of expression that tend to correlate with tempo. For example, using event loudness in the prediction would simply require the addition of another layer of variables in the Bayesian network and conditioning the timing variables on these nodes as well.

5.2 Capturing Style

Much work remains to expand on the characterization of stylistic mode. As previously mentioned, we plan to explore segmental stylistic characterization, considering different contextual information for each part of the performance. In our current model we use only one stylistic node. This may be a plausible for a small segment of music, but in a longer performance the choice of performance style may vary over time. If the predicted performance starts within one style but changes to another, the model is ill-informed to predict the parameters. In our future work we would like to extend the model to capture such stylistic tendencies over time. One approach would require pre-segmentation of the piece based on the choice of expressive choices during the reharsal stage, and introduction of one stylistic node per segment. The prediction context would then be local to each part of the performance. We may then, for example, have causal conditional dependencies between the stylistic nodes in each segment of the piece, which would allow the system to both infer the style within a part of the performance from what is being played and from the previous stylistic choices.

In practice, a musician or ensemble's rehearsals may not comprise of completely distinct interpretations; however, capturing expression contextually will likely offer a larger degree of freedom to the musicians in an internet performance, who may then explore a greater variety of temporal and other articulations.

5.3 Virtual Cueing

Virtual cueing forms an additional application of interest. As mentioned at the start of the paper, visual communication is generally absent or otherwise delayed in network music performance. If we could predict with reasonable accuracy the timing in sections of a piece requiring temporal coordination, then we could help musicians synchronize by providing them with perfectly simultaneous predicted cues. We regard the use of predictive virtual cues as less invasive to networked ensembles than complete predictive sonification. In situations where the audio latency is low enough for performance to be feasible but video latency is still too high for effective transmission of gestural cues, predictive sonification may be omitted completely, and virtual cues could be implemented as a regulating factor.

6. ACKNOWLEDGEMENTS

This research was funded in part by the Engineering and Physical Sciences Research Council.

7. REFERENCES

[1] C. Alexandraki and R. Bader. Using computer accompaniment to assist networked music performance. In *Proc. of the AES 53rd Conference on Semantic Audio, London, UK*, 2013.

[2] C. Chafe. Tapping into the internet as an acoustical/musical medium. *Contemporary Music Review*, 28, Issue 4:413–420, 2010.

[3] C. Chafe and M. Gurevich. Network time delay and ensemble accuracy: Effects of latency, asymmetry. In *Proc. of the 117th Audio Engineering Society Convention*, 2004.

[4] E. Chew and C. Callender. Conceptual and experiential representations of tempo: Effects on expressive performance comparisons. In *Proc. of the 4th International Conference on Mathematics and Compution in Music*, pages 76–87, 2013.

[5] E. Chew, A. Sawchuk, C. Tanoue, and R. Zimmermann. Segmental tempo analysis of performances in user-centered experiments in the distributed immersive performance project. In *Proc. of the Sound and Music Computing Conference*, 2005.

[6] E. Chew, R. Zimmermann, A. Sawchuk, C. Kyriakakis, and C. Papadopolous. Musical interaction at a distance: Distributed immersive performance. In *Proc. of the 4th Open Workshop of MUSICNETWORK, Barcelona*, 2004.

[7] A. Cont. Antescofo: Anticipatory synchronization and control of interactive parameters in computer music. In *Proc. of the International Computer Music Conference*, 2008.

[8] S. Flossmann, M. Grachten, and G. Widmer. *Guide to Computing for Expressive Music Performance*, chapter Expressive Performance Rendering with Probabilistic Models, pages 75–98. Springer Verlag, 2013.

[9] K. P. Murphy. Fitting a conditional linear gaussian distribution. Technical report, University of British Columbia, 1998.

[10] D. Pelleg and A. W. Moore. X-means: Extending k-means with efficient estimation of the number of clusters. In *Proc. of the Seventeenth International Conference on Machine Learning*, pages 727–734, 2000.

[11] C. Raphael. Music plus one and machine learning. In *Proc. of the 27th International Conference on Machine Learning*, pages 21–28, 2010.

[12] M. Sarkar. Tablanet: a real-time online musical collaboration system for indian percussion. Master's thesis, MIT, 2007.