ANALYZING CHICKEN GROUPS:

Here we explain the process followed by experts when analyzing the chicken groups. We start off by discussing the state variant sequence analysis of two groups, we follow this with an illustration of how the music notation is used to identify behaviors of subjects and then show how these behaviors affect ranking algorithms.

Using the state variant sequence and music notation to investigate a single group:

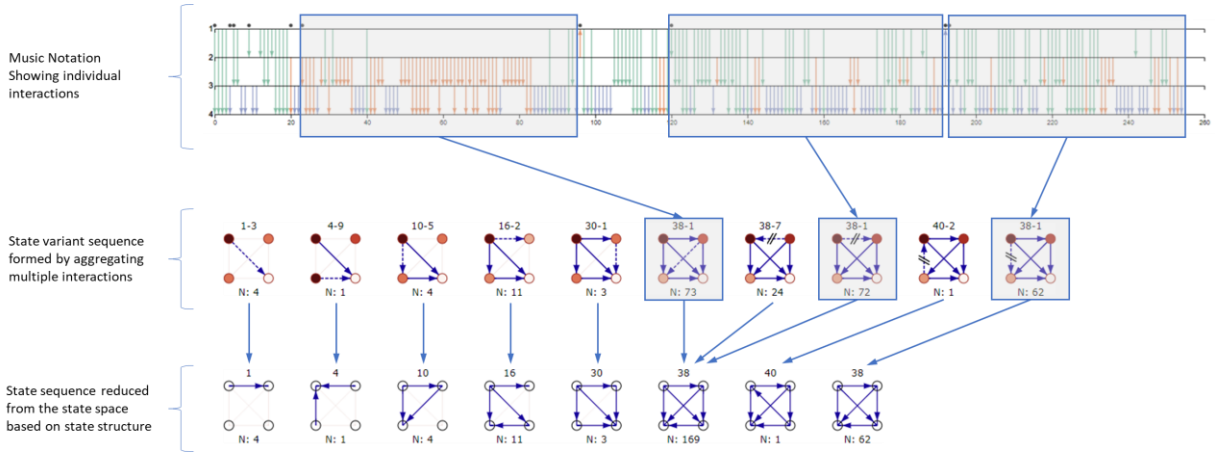

Figure 1. The individual interactions for group 1 shown in the music notation in the first row are aggregated and represented by the state variant sequence in the second row. The state variant sequence is further compressed by mapping each state into a unique configuration forming the state sequence.

The experts started their analyses by first examining the state variant sequences while using the state sequence and music notation to support their investigation of the hierarchies. Figure 1 shows how interactions from group 1 are represented and mapped to different representations. The two windows marked on the music notation map to two identical state variants (ID 38-1 labeled above the state and highlighted in the figure) which are also identical to the last variant in the sequence. These state variants are made up of a little over 80% the interactions in the group thus showing that members are settled into this configuration. Further investigation of this stable state variant shows that subject ranked 1 attacks all subjects ranked below it, ranked 2 and 3 do the same while the subject with the lowest rank does not commit any attacks at all. This is the most common type of hierarchy observed. The state variant 38-7 (fourth from the end of the sequence) also contains a significant amount of interactions, 24 interactions to be exact. In this variant we see that subject 2 retaliates against subject 1, however, we see that this retaliation only occurs once from the music notation. This indicates that the rest of the hierarchy was actually in the same state for the other 23 interactions.

Also, this group, subjects ranked 1, 2, and 3 forms an intransitive triad in the second to last state variant (ID 40-2). This state variant only lasts for one interaction which was a single retaliation by the third-ranked subject against the first. After this interaction, the first ranked subject goes back to dominating the third-ranked subject and the group settles back into the stable state variant. On the other hand, in group 5 we see in its state variant sequence in figure 2, we observe that there are a fair amount of reversals. On closer observation, we see that either subject ranked 1 or 3 are involved with the reversals. This prompted our experts to schedule a meeting among themselves to further investigate this phenomenon. Next, the experts wanted to look at cycles or intransitive triads, this is shown in figure 2 on the right. They observed that there are a high number of these cycles in this group as compared to other chicken groups. These cycles occurred at the start during rank contentions and did not persist for long durations (either 1 or 2 interactions per intransitive state). Thus, they concluded that these triads do not have a significant effect on the ranks.

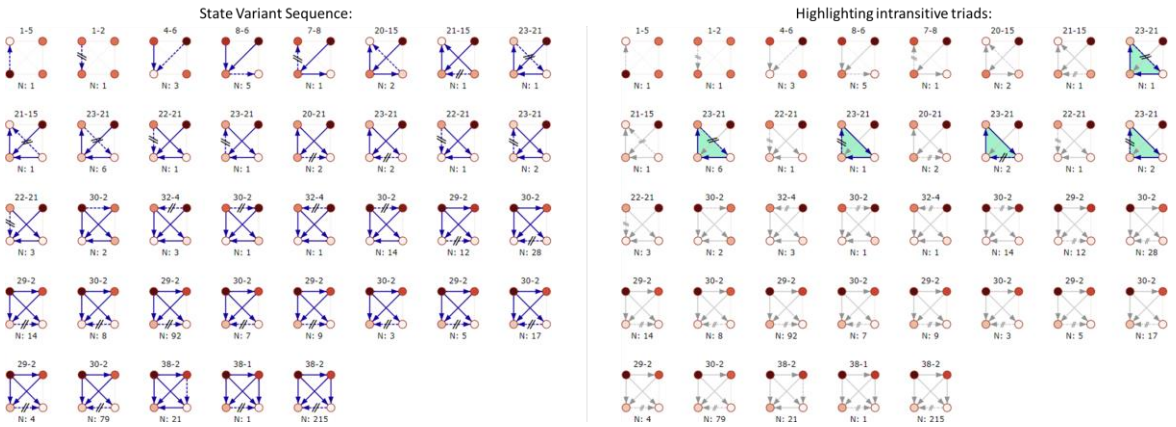

Figure 2. The state variant sequence *for group 5 on the left and the intransitive triads for the same group highlighted on the right.*

Using the music notation to investigate groups:

The experts also analyze the general behavior of subjects in a group through the music notation. When examining the different chicken groups, they looked at the music notations to identify any observable unique behaviors. We show three different music notations that were very different from each other in figure 3 and figure 4.

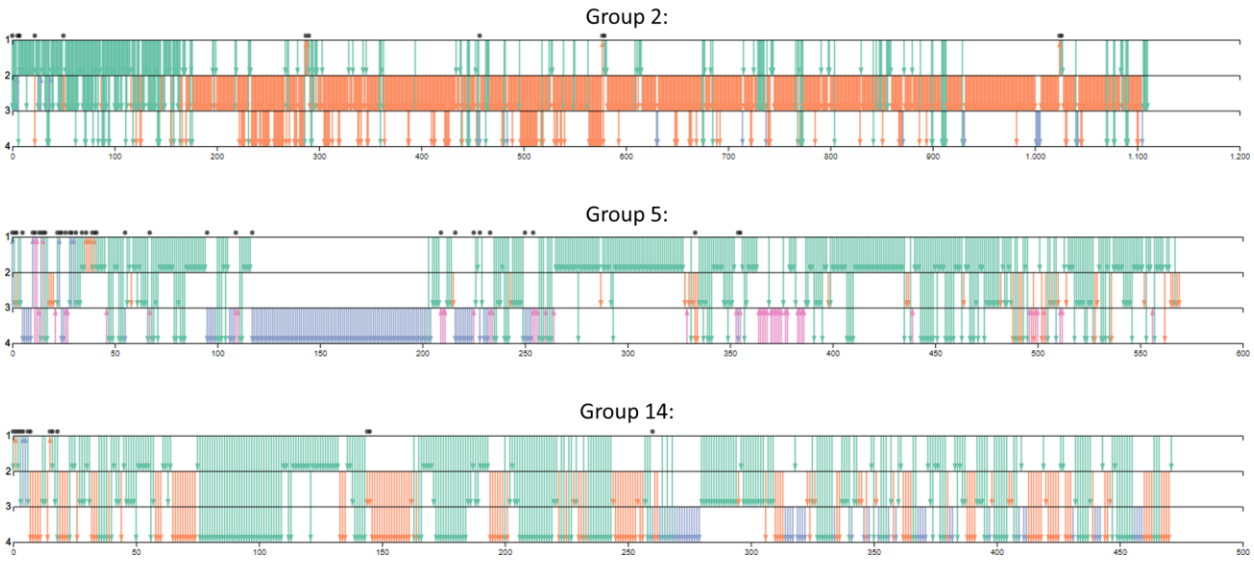

Figure 3. The music notations for group 2, 5, and 14. Group 2 shows two indivuals committing most attacks. Group 14 shows that the top ranked indiviual commits most. Group 5 is unusual, it shows that the lowest ranked subject is not as submissive as the lowest ranked indivual in other groups and it retaliates against the subject ranked just above it in the second half of the sequence.

Figure 3 shows three groups – 2, 5, and 14. Observing the music notations, we first see that group 2 has an unusually large amount of interactions (N = 1111), more than double the average number of interactions (N = 518) across 14 chicken groups. In group 2, we observe that most interactions were initiated by subjects ranked 1 and 2. Also, what is unusual is that subject 2 commits almost 3 times the number of attacks as subject 1, however it is ranked lower. This due to subject 1 mostly attacking subject 2 and subject 2 rarely retaliating. The distribution of interactions among members in Group 14 seems to be more like that of other groups, with the top-ranking subject initiating most of the attacks. However, at a later stage (in the MDS plot) it is shown to be different. Group 5 at first glance doesn't look odd, but on closer observation, we see that the lowest ranked subject attacks the 3rd ranked subject (pink arrows) a fair number of times in the second half of the music notation. It is also noticeable that the 3rd ranked subject doesn't retaliate. This is unusual behavior which affects ranking algorithms is further investigated in the following section with the ranking and balloon plot.

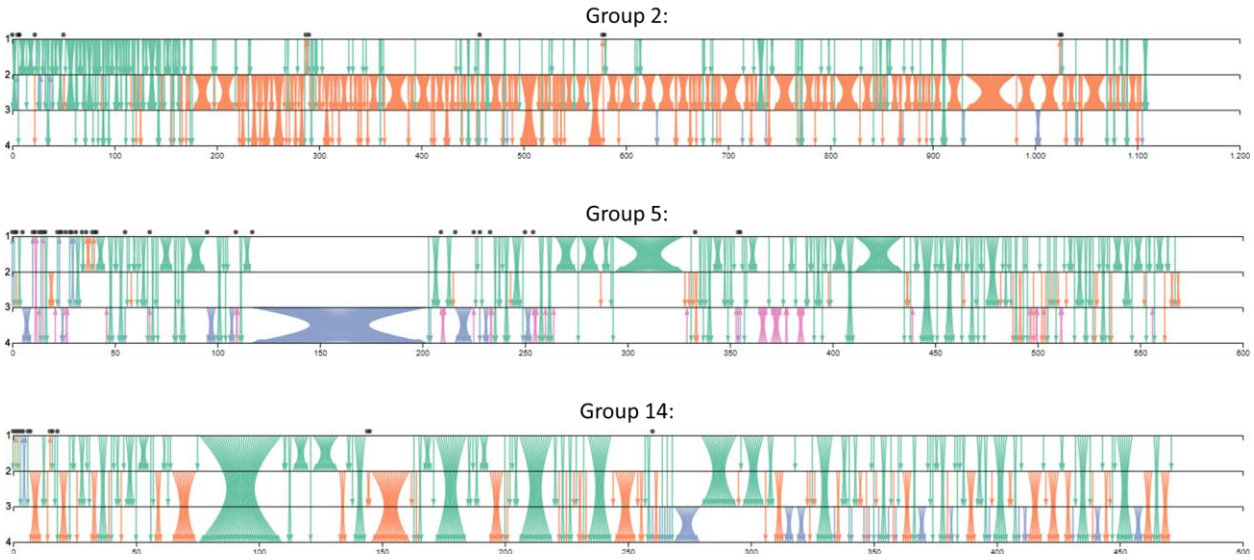

Figure 4. The music notations for group 2, 14 and 15 with bundling of bursts turned on. Group 2 shows that the subject ranked 2 continues to burst against the third ranked subject after the top ranked subject cements its place. Groups 5 and 14 show that the top ranked subject commits a few large bursts and multiple small bursts that reduce in size towards the end. Additionally, in group 14 also it is noticeable that the top ranked subject bursts against all other subjects but the second ranked subject almost never bursts against the third ranked subject this may indicate that the 2nd and 3rd ranked subjects may be closer in rank than computer by the ranking algorithm.

We also use the music notation to examine the bursting behaviors of individual subjects. To help users visually identify these burst clearly, we allow the user to turn on edge bundling for bursts as shown in figure 4. Here we see in Group 2 the subject ranked one first cements its place in the hierarchy initially with multiple small bursts against subject 2 and 3. Following this, the subject ranked 2 continues to burst against the third-ranked subject until the end of the interaction sequence. On the other hand, in group 5 and 14 commits relatively larger bursts and they occur throughout the interaction sequences. In both groups, the size of bursts reduces as time progresses. We also notice that in group 14 the top-ranked subject bursts against all other subjects but the second-ranked subject almost never bursts against the third-ranked subject this may indicate that the 2nd and 3rd ranked subjects may be closer in rank than computed by the ranking algorithm.

Using the ranking chart and balloon plot to investigate a group:

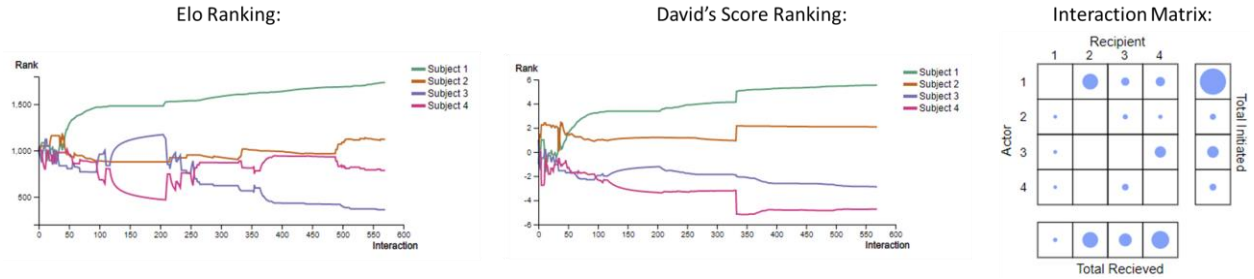

Figure 5. The Elo ranking and David's score ranking charts for group 5 showing different ranks for subjects 3 and 4 along with the interaction matrix showing the actual number of interactions with a balloon plot. Note that in this case the subjects were labeled (1 to 4) with the ranking generated with David's Score.

The Elo rankings, David's score rankings and the balloon chart for group 5 are shown in figure 5. As discussed, group 5 exhibits unusual behavior by the bottom two ranked subjects (3 and 4). We first examine the rank evolution chart with the different ranking algorithms, we observe that the Elo ranking and David's score interchanges the ranks of these two subjects. Additionally, David's score clearly separates them from the individuals ranked 1 and 2 at the very start unlike the Elo algorithm that shows a fair amount of competition among subjects 2, 3 and 4. Looking at the balloon chart, we observe that the two subjects 3 and 4 never attack the second-ranked subject but both of them perform a similar low number of attacks on the top-ranked subject. Going back to the music notation we found that these interactions occurred at the start of the sequence after which subject 1 retaliated and remained dominant. The balloon plot also shows that the 3rd ranked subject ranked 3 performs almost double the number of attacks against subject 4 as compared to the occurrence of the reverse, however, the Elo algorithm ranks 3 below 4.

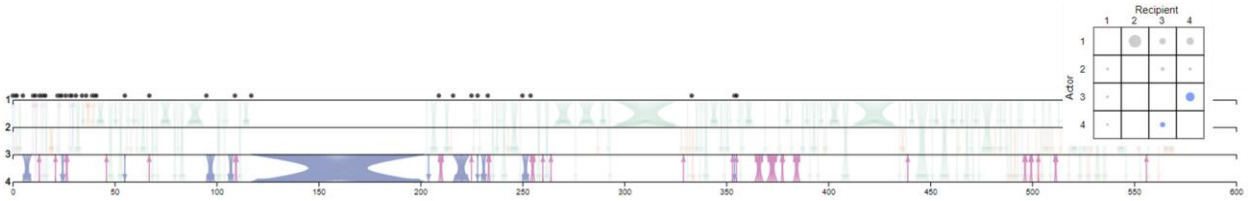

Figure 6. The filtered music notation for group 5 showing the interactions between subject 1 (green) and 3 (blue). And the balloon plot reflecting the selection (blue circles).

We investigate further by filtering these interactions and examining the music notation as shown in figure 6. We observe that subject 3 heavily attacks subject 4 at the start but as time progresses subject 3 retaliates and subject 4 becomes submissive. Thus, we hypothesize that Elo rankings are better for showing a more instantaneous rank while David's score is more reflective of the historical ranks. These ranking techniques can be used interchangeably to analyze the groups over a long period of time and it could possibly show how dominant members lose rank with age while younger members gain rank.

Investigating group similarities:

Using the techniques described above the experts analyzed the individual groups. Next, they wanted to compare these groups. They first compared the groups ranking charts represented as small multiples (figure 7). They immediately noticed that the chickens tend to form clear hierarchies early on.

Figure 7. Ranking charts for the 14 chicken groups (quads) represented as small multiples.

Next the experts looked at the chart in figure 8 which is an expanded version of the COF and CSF heatmap table and illustrates the CSF. It shows the configurations or states formed in each of the chicken groups. They observed that the chickens do not form box configurations (24-27) the chickens also do not form configuration 9 where one member is attacked by three others who have never been attacked. The experts also observed that state 38 was formed by all groups. It was difficult for the experts to use this chart on its own to determine how similar groups were, thus we provide an MDS plot as described in the paper that measured group similarity based on the DTW distance. The DTW distance was computed over the chain of unique configurations formed (an example is shown in figure 9) by each group. These distances were then represented with an MDS plot in figure 10.

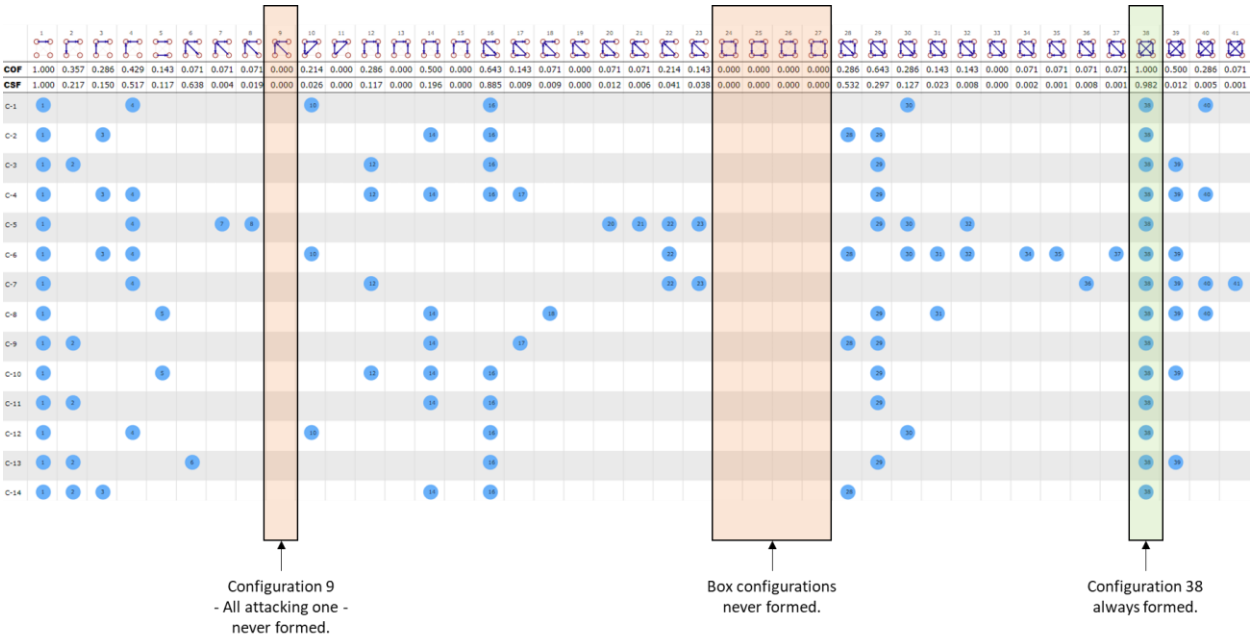

Figure 8. The chart showing the configurations formed by the 14 chicken groups. It is annotated with the expert's observations.

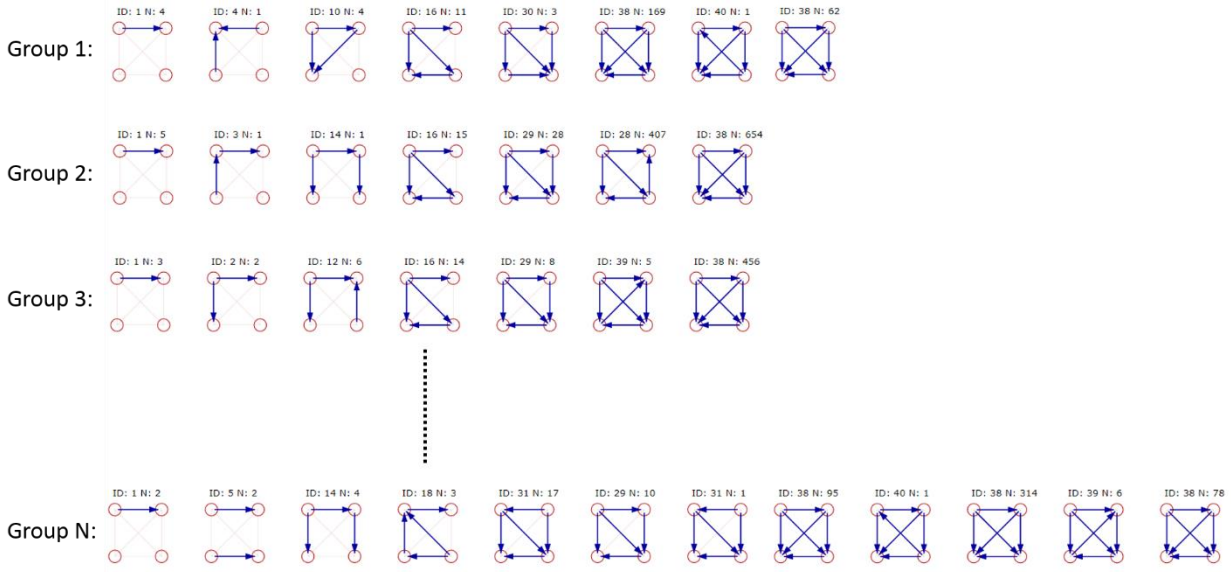

Figure 9. The states formed chains for the chicken groups. The chains vary in length but do have common sequences of configurations.

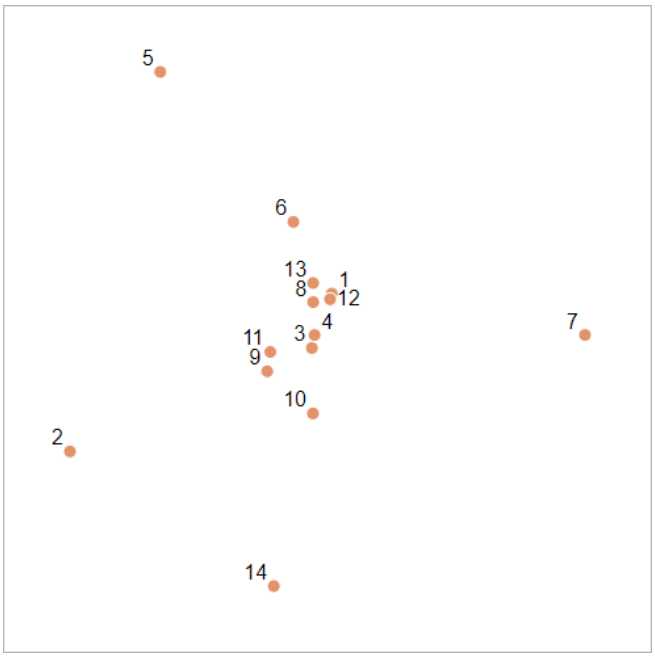

Figure 10. An MDS plot showing the similarity of groups based on a DTW distance. We see four outliers - group 2, 5, 7, and 14.

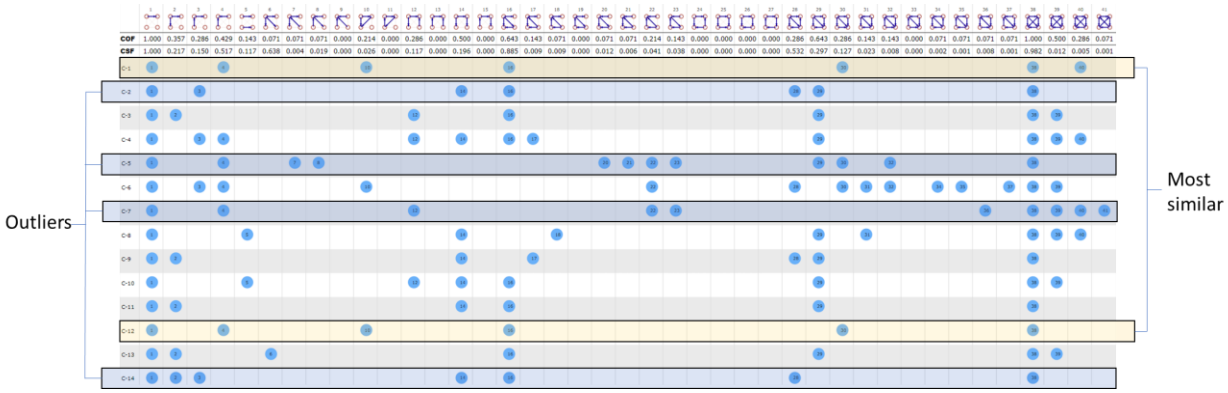

Figure 11. Two groups 1 and 12 that are very similar and four outliers from the MDS plot - groups 2, 5, 7, and 14 - are marked for reference.

Using the MDS plot (figure 10) and the configurations formed chart (figure 11). We see that group 1 and 12 are very similar, the main difference can be deduced by looking at the chart in figure 10 where we see that group 1 forms configuration 40 but not 2. It also stays in that configuration for a very short period (just 1 interaction) thus the DTW distance is very small. By looking at the outliers in figure 10 we see that they each form a different set of configurations that are not formed in other groups. For example, group 5 forms configurations 7, 8, 20 and 21 which are not formed in any other group. Similarly, group 7 is the only one to form configurations 36 and 41. However figure 11 does not indicate why group 2 and 14 were outliers, to investigate this the experts had to go back and analyze the state sequence where they observed an unusual ordering and stability of the configurations.

ANALYZING DEBATE DATA:

While our system was primarily designed to investigate hierarchies in animal groups, we can apply it to analyze dominance hierarchies in other situations just as well. One such situation is debates. To demonstrate this, we used our system to analyze aggressive behaviors in the 2016 U.S. presidential debates. Our system analyzes aggressive behaviors and for the purpose of this demonstration, we only categorized interrupts during debates as aggressive behavior. We used the transcribed debate data and analysis technique from Stephanie Kirmer's post on Kaggle (https://www.kaggle.com/skirmer/interruptions-at-the-first-presidential-debate) to extract these interruptions. With a more sophisticated analysis, one could extract aggressive responses by the candidates as well. For this analysis, we had two groups, the presidential debate with Donald Trump, Hilary Clinton, the moderator, and the audience as group members, and the vice-presidential debate with Mike Pence, Tim Kaine, the moderator, and the audience.

We approached the analysis in a manner similar to that followed by the experts as described in the paper. We first inspected the state sequence and noticed that in the presidential a complete six-link hierarchy (state) was formed but in the vice-presidential debate only a three-link hierarchy was formed. On further inspection were learned that the reason for this was that the audience commits an interruption (of Trump) in the presidential debate but they do not interact in the vice-presidential debate. The interruption is the 81st interaction in the music notation shown in figure 12 (top row). Next, we inspected the music notations, rank evolution and balloon plot of the two groups as in figure 12. Here we observed that in the debate with the presidential candidates Trump committed a burst of interrupts at multiple points in the debate. However, in the vice-presidential debates each candidate committed a burst of interrupts at the start of the debate after which the candidates and the moderator never committed more than two sequential interrupts. The ranking chart and the balloon plot informed us that in the presidential debate Trump committed more than half of the total number of interruptions and interrupted Clinton and the moderator equally. This also caused Trump to be ranked very highly during the debate but towards the end, the ranks of the group members started to converge as the moderator interrupted Trump more frequently toward the end. In the vice presidential the debate candidates interrupted each other equally and thus no single member was dominant for a long period of time. The state variant sequence for both debates was very long with very few stable states informing us that the groups never achieved a distinct hierarchy. This is reflective of highly competitive debates which these were. Finally, as we had just two groups, comparing them with the inter-group representations did not reveal any new information.

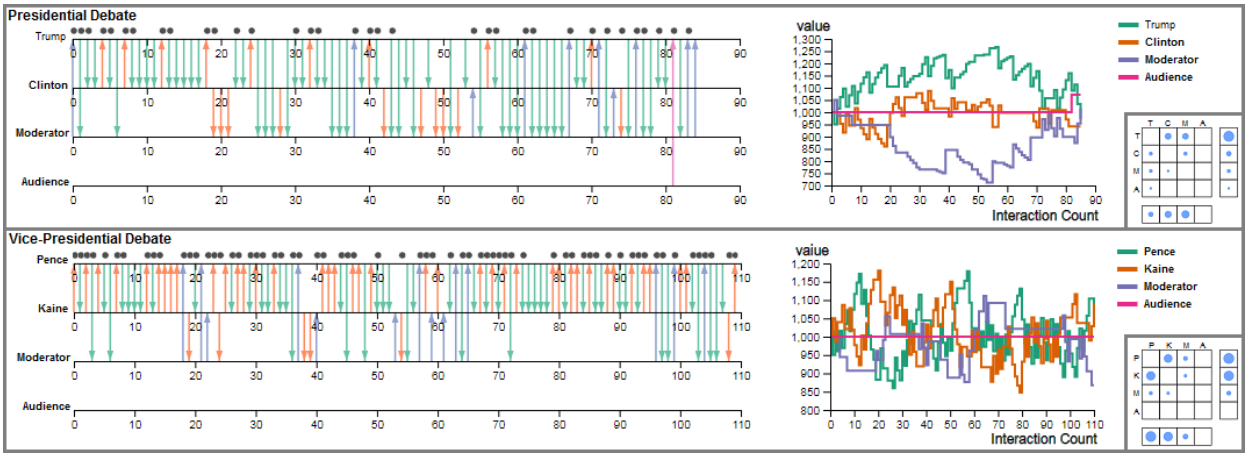

Figure 12. The music notation representation, Elo ranking chart, and the balloon plot for the presidential (top) and vice-presidential debates. The top row shows multiple bursts of interrupts from Trump but eventually the two debaters' ranks converged, it also shows he committed the most interrupts. But in the vice-presidential debate the candidates each had one burst of interrupts (between interactions 10-20 and 70-80). However, they both committed a similar number of interrupts and kept interrupting each other thus never creating a clear hierarchy.