A Markov Model for Volleyball

Brian Xu

Abstract

Current passing metrics are subjective and depend heavily on human judgement. We describe a Markov chain designed by Florence et al. to model the progression of a volleyball rally and our improvements to the model that enable us to create a better passing metric. Finally, we discuss the results after implementing our Markov model and determine which passes contribute the most to winning using our passing metric.

Department of Statistics

Rice University

August 2021

Contents

List of Figures

List of Tables

1 Introduction

In volleyball, current passing metrics are subjective. Passing is graded on a three point (0–3) scale based on the number of hitters the setter can set with the pass. For example, if the receiver gets aced or overpasses, the pass is rated as a zero since zero hitters are settable. If the receiver passes the ball such that the setter can only set one hitter, the pass is rated as a one and so on. However, besides the zero point rating, all the other ratings are prone to subjectivity. The line between whether or not a pass is good enough to set all three hitting options or only two is easily blurred by external factors such as ability of the setter and strictness of the passing grader. Graders also evaluate the location and height of a pass when determining a pass's rating. This dynamic creates inconsistency between pass ratings depending on the grader. Other rating systems for passing that use different point scales exist, but they also suffer from the same issues as the three point passing scale. We seek to create a better passing metric that eliminates human subjectivity.

We use a Markov chain to create a better passing metric. A Markov chain is a stochastic model that satisfies the following condition: the probability distribution of future states, conditional on the present and past values of the states, depends only on the present state. Knowing past states does not change the probability distribution of future states given that we know the present state. This condition is known as the Markov property. We use a Markov chain for 3 reasons. First, the progression of a rally in volleyball is easily modeled by a Markov chain. Each team is only allowed three touches before they have to send the ball over the net which results in offense following a rigid structure—pass, set, hit—most of the time. This structure can be clearly delineated as a sequence of states. Second, volleyball touches are not independent. A poor reception makes it harder for the setter to set a good set, and a poor set makes it harder for a hitter to kill the ball. However, suppose that we have a setter that makes the same set twice in a row, once off a bad pass and once off a good pass. A hitter will still have the same probability of killing the ball regardless of how good the pass was since both sets are identical. If we extend this logic to all touches of a volleyball rally, it is easy to see that the same condition holds for all touches. A setter has the same probability of making a specific set from the same pass regardless of how difficult the serve was. A defender has the same probability of making a specific dig from the same hit regardless of how good the set was. Thus, the Markov property is satisfied since knowing past states does not change the conditional distribution of future states when given the present state. Third, Markov models have been applied to other sports such as baseball with success [1].

2 Data

Data from official NCAA matches between the top 15 Divison 1 Women's Volleyball Teams during the 2019 season was collected from VolleyMetrics [2]. In total, 185 matches and 25,826 passes worth of data was recorded. Data from VolleyMetrics is recorded manually using Data Volley software which helps record data through the use of keyboard shortcuts. These Data Volley files are uploaded onto the VolleyMetrics site. Data Volley files contain data on each touch during a volleyball match. For every touch, analysts record the skill being performed, evaluation of the touch, start and end location of the touch, player performing the touch, team the player is on, which rally the touch occurred on, and more.

3 Model Design

Markov models have already been applied to volleyball. In 2008, a first-order Markov chain that modeled the progression of a volleyball rally was developed by Florence et al. The states of the Markov model consisted of the possible touches during a volleyball rally such as pass, set, hit, etcetera and the quality of the touch. A five point (0–4) rating system was used to grade passing. Sets were evaluated based on distance from net, height of set, and location of set relative to the hitter. There are three outcomes: rally won, rally lost, or rally continuation. Thus a possible sequence of states could be four-point pass to low and inside set 0–3 feet from net to rally won [3]. This sequence models a rally where a receiver perfectly passes the ball, then the setter sets the ball low and inside and tight to the net, and finally the hitter kills the ball for a point. This model captures the progression of a volleyball rally well, however, it uses passing metrics that are susceptible to subjectivity for reasons discussed in the introduction. We seek to expand upon this model by using positional data for passing instead.

Our model is based on the Markov chain developed by Florence et al. The states of our first-order Markov chain consist of the possible touches during an offensive possession of a volleyball rally and the quality of the touch. However, instead of using a rating system to grade passing, we categorize passes by their end location, or the location the setter receives the pass in. The volleyball court is split into nine zones, where each zone is composed of four sub-zones for a total of 36 positions. Figure 1 shows the position of each zone. The thick black line represents the net. Figure 2 shows the position of each sub-zone within a zone. We also include passing error and overpass as states, making the total number of passing states 38. Our data does not contain the same setting metrics used by Florence et al. Instead, we use the set ratings given by the Data Volley graders. There are three ratings for sets: perfect, poor, error. Since set errors are usually the fault of setters, and they occur infrequently (only 1.5% of sets were rated errors), we will exclude set error as a state. Our three outcome states for each rally are rally won, rally lost, or rally continuation. Finally, our starting state is always the opponent serve since volleyball rallies always start with a serve.

4	3	2
7	8	9
5	6	1

Figure 1: Zones of a volleyball court

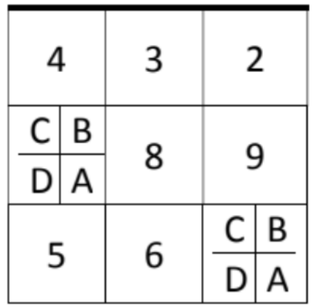

Figure 2: Sub-zones of a volleyball court

The model is from the perspective of the receiving team, so rally won means the receiving team won the rally and rally lost means the serving team won the rally. We can think of our Markov model as a tuple containing Q, A, and π where Q is the set of states, A is the transition matrix, and π is the starting state vector. We have a total of 44 states—one starting state for the serve, thirty-six possible passes, two possible sets, and three rally outcomes. Thus, our transition matrix will have 44 rows and columns to represent the transition probability between each state. Since A is a stochastic matrix, the sum of all elements in each row is one. π is a vector of length 44 where $\pi_{serve} = 1$ and $\pi_i = 0$ for all $i \in Q \setminus \{serve\}$ since a volleyball rally always starts with a serve.

Our model has made a number of simplifications. We assume that setters will never perform an illegal action such as doubling or carrying that results in a setter error and assume that setters will never attack the ball on the second touch. Both these actions occur in volleyball games, but we exclude them from our model to represent the offensive structure of pass, set, hit that the vast majority of volleyball rallies follow. We have left out the hitting as a state, because our three outcome states already account for the hitting touch. The hitter can either spike the ball such that they score, spike the ball such that their team loses the point, or spike the ball such that the opponents dig the ball or their team recovers the ball from a block, resulting in rally won, rally lost, and rally continuation respectively.

Now that we have developed the Markov model, we can discuss how the model will help us create a better passing metric. We can calculate the probability of the receiving team winning the rally on the first possession given that they pass the ball. This probability is known as First Ball Side-out Percentage (FBSO%).

$$\text{FBSO\%} = \frac{\text{First Ball Kills}}{\text{Receptions}} \cdot 100 \tag{1}$$

We will compare FBSO% to the probability of winning the rally given passes to specific locations. For example, if the probability of winning the rally given a pass to zone 3A $\big(P(W|3A)\big)$ is higher than (FBSO%), then the pass has increased our chance of winning the rally by $P(W|3A) - P(FBSO)$. We can extend this framework to every zone on the court, enabling us to determine how much a certain pass contributes to a team's chance of winning the rally. A player's average pass rating for a match would then be

$$\text{Pass Rating} = \frac{\big(\sum_{\forall i \in P} i\big) - |P|(FBSO)}{|P|} = \frac{\big(\sum_{\forall i \in P} i\big)}{|P|} - FBSO \tag{2}$$

where P is an array of $P(W|X) - P(FBSO)$ values for each pass the passer has made, $P(W)$ is the probability of winning the rally, X is an element from the set of pass states in our Markov model, and $FBSO$ is the probability of the receiving team winning the rally on the first possession after a pass has been made.

4 Model Implementation

All analysis was done using R [4]. The R package, datavolley, was used to read in the data from Data Volley files into a data frame [5]. Transition probabilities between states were calculated using point estimation by observing the number of transitions between each state. We define the following counts:

- $C(s_i)$: The number of times state s_i appears.
- $C(s_{i-1}, s_i)$: The number of times the state sequence s_{i-1}, s_i appears.

Then, the transition probability between state s_{i-1} and state s_i is

$$P(s_i|s_{i-1}) = \frac{C(s_{i-1}, s_i)}{C(s_{i-1})} \tag{3}$$

To implement the model, we created a sequence of states of every first ball volleyball touch in the data frame. Then, we iterated over the sequence of states to compute $C(s_i)$ and $C(s_{i-1}, s_i)$ for every state in Q. Finally, we calculated the transition probabilities between all states using Equation 3. Impossible state transitions such as pass to serve were constrained to a probability of zero. Table 1 shows the transition matrix with transition probabilities calculated.

Surprisingly, poor sets have a higher probability of resulting in a winning attack compared to perfect sets. This result is not meaningful however, since only 610 sets were rated poor compared to the 22,436 sets that were rated perfect.

	Serve	Pass Error	Overpass	P1A	P1B	P1C	P1D	P2A	P2B	P2C	P2D	P3A	P3B	P3C	P3D	P4A	P4B	P4C	P4D	P5A	P5B	P5C	P5D	P6A	P6B	P6C	P6D	P7A	P7B	P7C	P7D	P8A	P8B	P8C	P8D	P9A	P9B	P9C	P9D	Perfect Set	Poor Set	Rally Won	Rally Lost	Rally Continuation
Serve	0.000	0.055	0.038	0.001	0.002	0.004	0.001	0.015	0.008	0.074	0.110	0.144	0.106	0.066	0.082	0.019	0.013	0.002	0.004	0.001	0.003	0.001	0.001	0.001	0.004	0.004	0.000	0.005	0.010	0.003	0.002	0.013	0.036	0.023	0.009	0.004	0.007	0.024	0.010	0.000	0.000	0.095	0.000	0.000
Pass Error	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
Overpass	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
P1A	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.250	0.750	0.000	0.000	0.000
P1B	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.786	0.214	0.000	0.000	0.000
P1C	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.906	0.094	0.000	0.000	0.000
P1D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.588	0.412	0.000	0.000	0.000
P2A	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.879	0.121	0.000	0.000	0.000
P2B	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.905	0.095	0.000	0.000	0.000
P2C	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.979	0.021	0.000	0.000	0.000
P2D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.986	0.014	0.000	0.000	0.000
P3A	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.990	0.010	0.000	0.000	0.000
P3B	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.983	0.017	0.000	0.000	0.000
P3C	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.978	0.022	0.000	0.000	0.000
P3D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.988	0.012	0.000	0.000	0.000
P4A	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.967	0.033	0.000	0.000	0.000
P4B	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.953	0.047	0.000	0.000	0.000
P4C	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.796	0.204	0.000	0.000	0.000
P4D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.730	0.270	0.000	0.000	0.000
P5A	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.842	0.158	0.000	0.000	0.000
P5B	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.924	0.076	0.000	0.000	0.000
P5C	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.718	0.282	0.000	0.000	0.000
P5D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.294	0.706	0.000	0.000	0.000
P6A	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.733	0.267	0.000	0.000	0.000
P6B	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.938	0.062	0.000	0.000	0.000
P6C	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.922	0.078	0.000	0.000	0.000
P6D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.778	0.222	0.000	0.000	0.000
P7A	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.938	0.062	0.000	0.000	0.000
P7B	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.965	0.035	0.000	0.000	0.000
P7C	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.716	0.284	0.000	0.000	0.000
P7D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.716	0.284	0.000	0.000	0.000
P8A	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.970	0.030	0.000	0.000	0.000
P8B	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.983	0.017	0.000	0.000	0.000
P8C	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.977	0.023	0.000	0.000	0.000
P8D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.971	0.029	0.000	0.000	0.000
P9A	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.738	0.262	0.000	0.000	0.000
P9B	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.854	0.146	0.000	0.000	0.000
P9C	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.970	0.030	0.000	0.000	0.000
P9D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.925	0.075	0.000	0.000	0.000
Perfect Set	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.378	0.155	0.467
Poor Set	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.409	0.153	0.439
Rally Won	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Rally Lost	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Rally Continuation	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 1: Transition Matrix

Furthermore, the distinction between perfect and poor sets is unclear. A[Serve, Rally Won] represents the probability of the server making a serve error and A[Serve, Pass Error] represents the probability of the server acing the receiving team. FBSO% was determined using the counts computed for each state. Since FBSO% is the probability of the receiving team winning the rally on the first possession given that they pass the ball, FBSO% = $\frac{\text{Rallies Won - Aces}}{\text{Receptions}} \cdot 100$. Using our data, we computed this number to be 38.1%. Source code for model implementation can be found at my GitHub: `https://github.com/bzx24/markov-volleyball`.

5 Results

We summarize results by evaluating the probabilities of moving from a certain pass to a rally outcome. We also use our pass rating framework to determine the best passing positions. These results are summarized in the following tables and figures.

Pass Type	Rally Won	Rally Lost	Rally Continuation
Pass Error	0.000	1.000	0.000
Overpass	0.000	0.000	1.000
P1A	0.400	0.200	0.400
P1B	0.268	0.071	0.661
P1C	0.198	0.207	0.595

P1D	0.312	0.000	0.688
P2A	0.257	0.179	0.564
P2B	0.286	0.203	0.512
P2C	0.406	0.153	0.441
P2D	0.397	0.152	0.451
P3A	0.418	0.147	0.435
P3B	0.442	0.145	0.414
P3C	0.417	0.151	0.433
P3D	0.365	0.157	0.478
P4A	0.296	0.176	0.528
P4B	0.330	0.130	0.540
P4C	0.245	0.189	0.566
P4D	0.316	0.163	0.520
P5A	0.105	0.211	0.684
P5B	0.207	0.185	0.609
P5C	0.184	0.053	0.763
P5D	0.412	0.118	0.471
P6A	0.200	0.200	0.600
P6B	0.279	0.162	0.559
P6C	0.262	0.194	0.544
P6D	0.222	0.111	0.667
P7A	0.336	0.133	0.531
P7B	0.294	0.184	0.521

P7C	0.319	0.097	0.583
P7D	0.242	0.152	0.606
P8A	0.279	0.214	0.507
P8B	0.339	0.155	0.505
P8C	0.333	0.169	0.498
P8D	0.303	0.144	0.554
P9A	0.228	0.208	0.564
P9B	0.279	0.206	0.515
P9C	0.300	0.164	0.537
P9D	0.320	0.163	0.517

Table 2: Probability Point Estimates for Pass Types

Pass Type	Pass Rating
Pass Error	-0.381
Overpass	-0.381
P1A	0.019
P1B	-0.114
P1C	-0.183
P1D	-0.069
P2A	-0.125
P2B	-0.096
P2C	0.025

P2D	0.016
P3A	0.036
P3B	0.060
P3C	0.035
P3D	-0.017
P4A	-0.085
P4B	-0.052
P4C	-0.136
P4D	-0.065
P5A	-0.276
P5B	-0.175
P5C	-0.197
P5D	0.030
P6A	-0.181
P6B	-0.102
P6C	-0.119
P6D	-0.159
P7A	-0.046
P7B	-0.087
P7C	-0.062
P7D	-0.139
P8A	-0.102
P8B	-0.042

P8C	-0.049
P8D	-0.079
P9A	-0.154
P9B	-0.102
P9C	-0.082
P9D	-0.062

Table 3: Pass Ratings for Pass Type

Pass Type	Pass Rating
P3B	0.060
P3A	0.036
P3C	0.035
P5D	0.030
P2C	0.025
P1A	0.019
P2D	0.016
P3D	-0.017
P8B	-0.042
P7A	-0.046
P8C	-0.049
P4B	-0.052
P9D	-0.062

P7C	-0.062
P4D	-0.065
P1D	-0.069
P8D	-0.079
P9C	-0.082
P4A	-0.085
P7B	-0.087
P2B	-0.096
P8A	-0.102
P9B	-0.102
P6B	-0.102
P1B	-0.114
P6C	-0.119
P2A	-0.125
P4C	-0.136
P7D	-0.139
P9A	-0.154
P6D	-0.159
P5B	-0.175
P6A	-0.181
P1C	-0.183
P5C	-0.197
P5A	-0.276

Pass Error	-0.381
Overpass	-0.381

Table 4: Pass Ratings for Pass Type Ordered

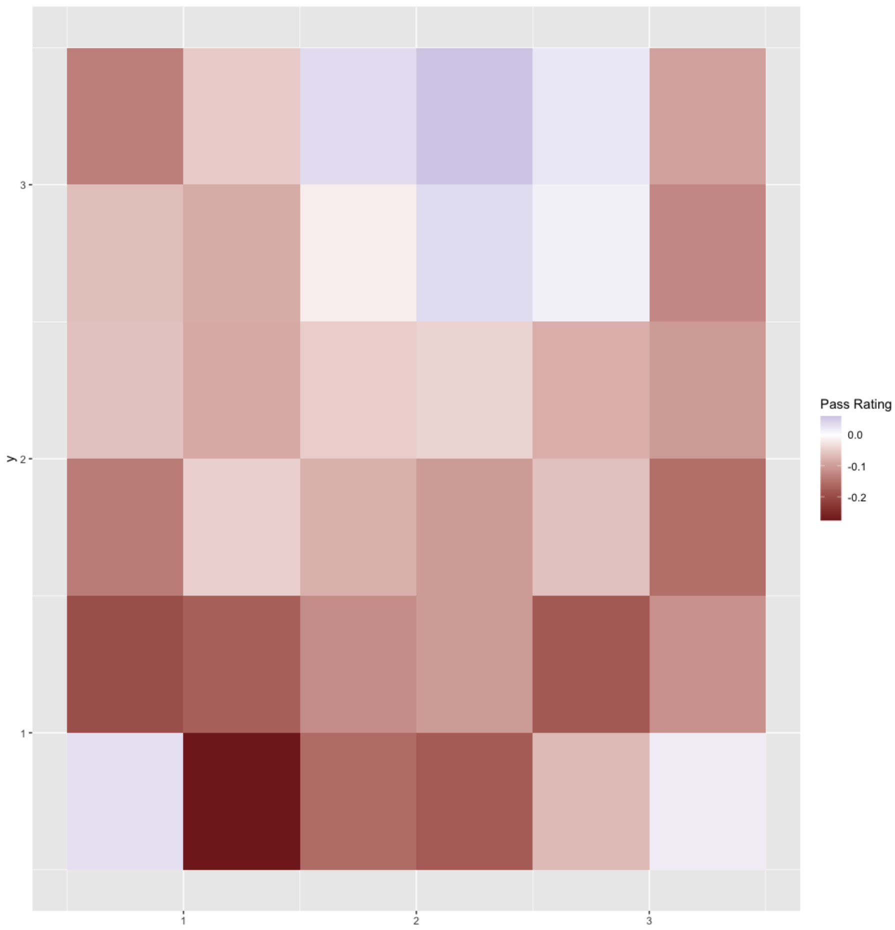

Figure 3: Pass Rating by Pass Zone

From Table 4, we see that the best spots to pass to are 3B, 3A, 3C, 5D, and 2C. Zone 3 and 2 are not a surprise since these zones are close to where the setter start at. However, zone 5 is the back left of the court in the complete opposite direction of the setter. This result is not meaningful though as only 17 balls were passed to zone 5D. Figure 3 shows which pass zones contribute to winning. Zones that are shades of blue represent pass zones that have a higher probability of siding out compared to FBSO% while zones that are shades of red represent pass zones that have a lower probability of siding out compared to FBSO%. The net is located at the top of the figure. From the figure shown, zones closer to the net centered in the middle of the court have the highest probability of resulting in a winning attack.

6 Conclusion and Future Work

We improve on the model developed by Florence et al. to create a better passing metric. Our metric determines how much a player's pass contributes to winning by comparing the probability of winning the rally given that specific pass to FBSO%. This rating system better evaluates the value of a specific pass because its based on win contribution, and ratings are not limited to a scale from 0–3 which allows for more nuanced distinction between different passes. Our metric also does not depend on the human subjectivity which current passing metrics fall victim to. Thus, our passing metric is preferable to current passing metrics. Adopting our passing metric enables coaches to determine the passers that contribute the most to winning.

In this work, we only modelled the first offensive possession of a volleyball rally. The Markov chain could be extended in a future project to include every single touch of a volleyball rally which would enable evaluation of skill importance. By using the framework we developed, we could determine how each touch affects the probability of winning the rally. The difference between the probability of winning the rally given a touch and the probability of winning the rally given the previous touch is the value the touch adds to the team's chances of winning. Extension of the framework to every touch in volleyball would enable us to evaluate which skills are most important to winning in volleyball and would enable us to create a player evaluation metric such as Wins Above Replacemnt (WAR) for baseball or Adjusted Plus-Minus (APM) that would measure a player's effectiveness and contribution towards winning.

Our setting data was also constrained to only two ratings: perfect and poor. Two ratings is not adequate to evaluate setting accurately, especially considering the vast majority of sets were rated perfect. In the future, locational set data such as the setting metrics used by Florence et al. could be collected to better improve the model. Likewise, height and angle data for passes would also improve the model by further categorizing different types of passes. More match data could be collected since some pass zones were passed to less than thirty times.

7 Volleyball Terminology

1. Point: A team wins a point when the ball touches the opponent's court or the opposing side performs an illegal action. Examples of illegal actions include being the last team to touch a ball before it goes out of bounds, not getting the ball over the net in three touches or less, or carrying the ball.

2. Rally: A rally begins when the ball is served into play and ends when a team earns a point. Each team can only touch the ball three times in succession and no player can touch the ball two times in a row.

3. Serve: To begin each rally, the team that won the previous point must serve the ball at the receiving team. The server must hit the ball to the other side so that it crosses the net and lands in bounds. The server must also not cross the end line during the serve.

4. Serve Error: When the server serves the ball into the net, out of bounds, or commits an illegal action, resulting in the opposing team getting a point.

5. Ace: A serve that scores immediately off passer error.

6. Pass: The first touch by the receiving team after the serve. Also known as a reception or bump.

7. Overpass: A pass where the ball goes over the net so that the opposing team gets possession of the ball.

8. Set: A touch where the intention of the touch is to pass the ball to a spiker so that they can attack it. Usually performed by the setter although anyone is allowed to set.

9. Attack: The player touches the ball with the intention of sending the ball across the net to score a point. Also known as a spike or hit.

10. Kill: An attack that immediately results in a point.

11. Attack Error: When the attacker hits the ball out of bounds, into the net, gets blocked, or commits an illegal action, resulting in the opposing team getting a point.

12. Block: Players can jump and try to block an opposing spiker's attack. A block that results in an immediate point is referred to as a stuff block.

13. Dig: Similar to a pass except it occurs when receiving an opposing attack.

References

[1] Ursin, Daniel Joseph. “A Markov Model for Baseball with Applications” (2014). Theses and Dissertations. 964.

[2] VolleyMetrics. Hudl. https://www.hudl.com/products/volleymetrics.

[3] Florence, Lindsay W. Fellingham, Gilbert W. Vehrs, Pat R. and Mortensen, Nina P. “Skill Evaluation in Women’s Volleyball,” Journal of Quantitative Analysis in Sports: Vol. 4: Iss. 2, Article 14.

[4] R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

[5] Raymond, Ben. datavolley: R package for reading volleyball scouting files. URL: https://github.com/openvolley/datavolley

[5] Trinsey, Joe. Assistant Coach, Canadian Women’s National Team.