Developing State-Based Recommendation Systems for Golf Training

(Technical Paper)

Distance Gains and Growing Pains: Addressing Distance Gains on the PGA TOUR

(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Christopher Kaylor
Spring, 2020

Technical Project Team Members
Kelly Rohrer
Jacob Ziller
Orlando Jimenez
Alanna Flores

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Prospectus

Introduction:

When the average person thinks of golf, they envision sweater vests, country clubs, and other antiquated trappings of the elite. However, modern golf is pushing the boundaries of this reputation. Just as analytics have pervaded other sports, golf is embracing the powerful insights of data-driven decision-making. Coaches and players are learning how data informs practice routines and playing strategies. Additionally, the modern golfer now fits an athletic and powerful mold. Several inputs – improved equipment, increased emphasis on fitness, and better knowledge of swing biomechanics – have amplified the abilities of professional and amateur golfers, especially the distance they can hit a golf ball. As the golf world processes these transitions, it has become clear it will take expertise and wisdom to sustain and grow the sport – a primary focus of this research project.

The technical portion of this proposal focuses on helping GameForge Golf, a start-up in the golf data industry, navigate data integration and statistical modeling. The company has amassed ample data concerning both practice routines and tournament scoring. GameForge wishes to find statistical links between the two, allowing them to recommend practice drills based on scoring, but several data imperfections must be addressed beforehand. The STS component of this proposal discusses the proliferation of distance on the PGA Tour. While distance gains over the past two decades have helped golf gain a more athletic perception, industry experts are concerned distance has made players one-dimensional and rendered classic courses obsolete. There is no consensus on how to address this issue, but an effective response is crucial to the longevity of the game. The technical portion of this thesis will detail the creation of

a data-driven practice recommendation system, and the STS section will explore how golf authorities should handle distance gains on the PGA Tour.

Technical Topic

Background:

In recent years, advanced analytics have revolutionized professional sports. General managers in Major League Baseball have embraced Sabermetrics, a phenomenon documented by the acclaimed book and film *Moneyball* (Lewis, 2003) (Miller, 2012). Analytics gurus like Houston Rockets general manager Daryl Morey have transformed the National Basketball Association by prioritizing three-point shooting (Kram, 2019). In the National Football League, coaches are rethinking conventional decision-making on 4th down plays because of statistical findings of front-office analysts (NYT 4th Down Bot, 2014). However, golf has been left behind by the analytics revolution. The habits of elite players are largely rooted in traditional wisdom. GameForge Golf is a golf data company seeking to modernize golf's usage of analytics. The company aims to help players and coaches build practice regimens by creating a recommendation system linking practice drills to on-course statistics. However, GameForge struggles with the scope of their data and the complexity of modeling this link.

Data Challenges:

The GameForge data is split into two components. The first is data detailing how individual players scored on roughly 2,700 unique drills. This data, ranging from 2017 to 2019, presents several challenges. First, individual drills are scored in many different ways, so comparing drill performance is tricky. Second, players sometimes go long stretches without logging drills, and tend to under-enter drills they have performed poorly on. The second component is scoring data. This set features over 20,000 rounds, logged between 2014 and 2019,

and describes players' scores and their performance in 28 statistical categories. Unfortunately, many rounds have impermissible or missing data, and filtering such entries shrinks the data set. Balancing data quality and data size is an important consideration which will require the use of tools like Microsoft Excel and RStudio.

Creation of Recommendation System:

Once these problems have been addressed, the overall challenge of creating a recommendation system linking drill and round data can be assessed. Connecting these two features presents several technical complications. First, different types of players need to be accounted for. In golf, players can use varying skillsets to score similarly. Players most likely benefit more from drills addressing weaknesses, so the system must account for differing strengths, weaknesses, and styles of players. Figures 1 & 2 illustrate different statistical makeups of 169 total rounds by two separate groups of players who all shot 2-under. These charts show that some players excel at ball-striking, while others thrive around the greens.

Per Rnd Stats	**Cluster 1**	**Cluster 2**
P6	**3.44**	2.05
Chip Ins	**0.14**	0.04
Total Putts	**16**	29.7
One Putts	**8.53**	6.26
Bombs	**0.63**	0.25

Figure 1: Greenside play comparisons of -2 rounds, found through K-means clustering (Kaylor, 2019)

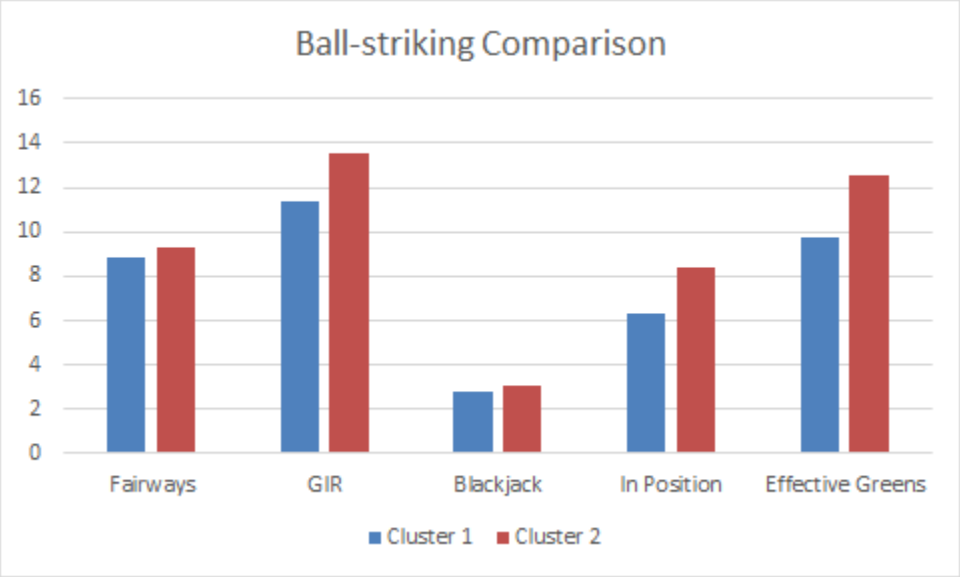
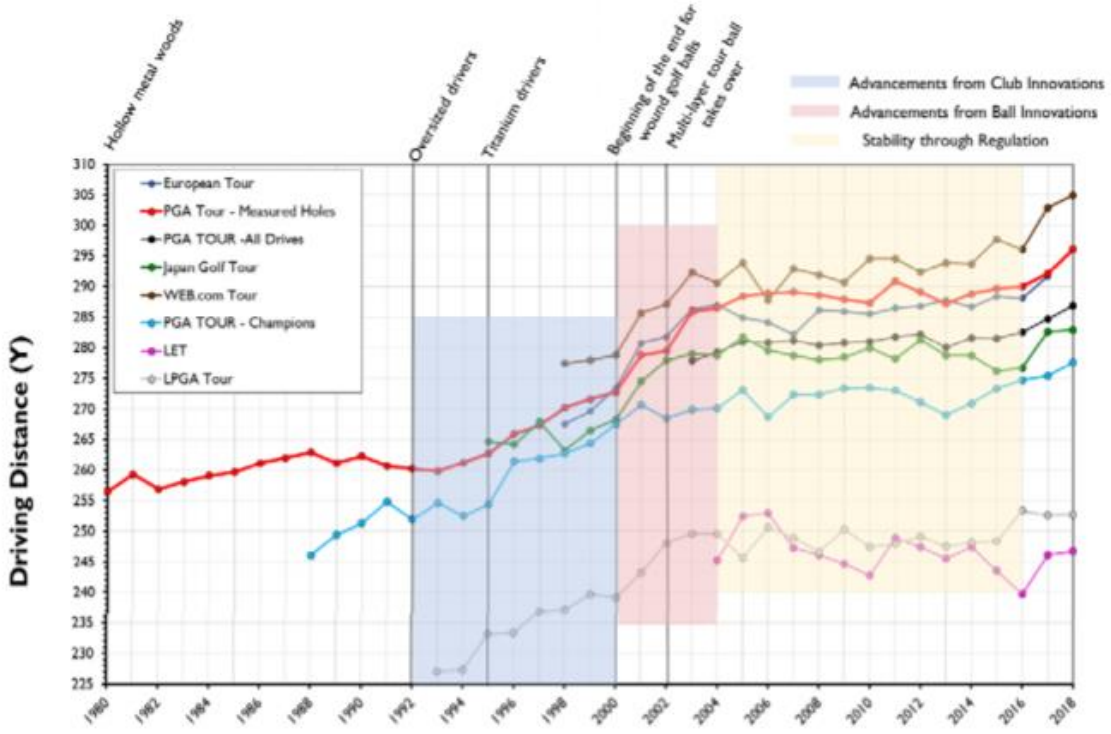


Figure 2: Ball-striking comparisons of -2 rounds, found through K-means clustering (Kaylor, 2019)

Second, it is uncertain if all drills performed by a player are equally significant to round performance. Golf courses, like players, vary in the skillsets they require, so drills focusing on course-specific skills could be more effective. Finally, making quantitative sense of the time between drills and rounds is complex. A player could perform the same drill twice, displaying equal skill each time, but the drill performed earlier might have less influence on an upcoming round. Accurately discounting the effect of less recent drills on round scoring will require quantitative modeling and contextual assumptions. In summary, the system must have the flexibility to account for variations in players, courses and time, while overcoming the limitations of the provided data. These technical findings will be completed by the end of the Fall 2019 semester, while a formal paper and presentation will be submitted to the SIEDS conference in May of 2020.

STS Topic

Background:



"*Figure 3.* Average driving distance on the major tours with significant innovation milestones overlaid" (*A Review of Driving Distance—2018*, 2018)

A major problem facing the PGA Tour over the past two decades is the dramatic increase in the distance players can hit the ball. This development has caused many to argue that players' skillsets are growing one-dimensional, classic course designs are falling into irrelevancy, and the PGA Tour is becoming a less compelling and relatable product. This distance phenomenon has gone largely undiscussed until recently – the United States Golf Association (USGA) only began publishing its annual distance report in 2015 (*A Review of Driving Distance—2018*, 2018). Some of this lack of awareness stems from ambiguity and multiplicity surrounding the causes of distance gains. The first two reasons, illustrated by Figure 3, both center around equipment. First, the multi-layered golf ball, pioneered by Titleist in 2000, caused a six-yard jump in average

driving distance in 2001 and has been improved upon since (McClusky, 2014). Second, improvements in driver materials and customizable features have helped players maximize both distance and accuracy (Rapaport, 2019b). Third, Tiger Woods' explosive play in the early 2000's inspired a younger generation of golfers to prioritize fitness (Kalland, 2016). Finally, swing instructors have used biomechanics to better understand how the golf swing creates stability and speed (Ritter, 2018). Increased distance on the PGA Tour is a complex phenomenon that cannot be attributed to a singular cause.

Stakeholders, Artifacts, and Perspectives:

This development has undeniably changed the complexion of PGA Tour golf. The strategies professional golfers use to attack classical courses has changed drastically, with one writer suggesting most courses "have been stretched to a near-breaking point" (Diaz, 2017) by record-breaking drives. This development is troublesome to the long-term appeal of the professional game, as many traditional fans worry creativity and strategy are becoming less integral parts of the game as distance becomes paramount. Former top-ranked players like Tiger Woods and Adam Scott have similar concerns, describing length as an increasingly meaningless design element of PGA Tour courses (Hoggard, 2017) (Marksbury, 2019). The three main governing bodies of golf have started to recognize the importance of addressing this issue. USGA president Mike Davis believes distance increases are "not necessarily good for the game" (Gray, 2018a) and Royal & Ancient chief Martin Slumbers suggests distance gains have crossed a "line in the sand" (Gray, 2018b). The third organization, the PGA Tour, has intensified its stance, already implementing a driver-testing policy designed to identify clubs that have slipped past permitted technological limits (Menta, 2019). Many independent voices in the game, such as all-time great player Jack Nicklaus, have suggested imposing design requirements on the golf

ball to limit distance (Mell, 2018). Others, such as current top player Brooks Koepka and ball manufacturer Taylormade, are not convinced this ball rollback would be a legitimate solution (Dethier, 2018) (Rapaport, 2019a). Despite the recency of the issue, distance gains on the PGA Tour have caught the attention of many important stakeholders in the game.

STS Framework:

With such a wide array of participants – players, fans, governing bodies, equipment companies, and golf course experts all come to mind – the use of Actor Network Theory (ANT) as an STS Framework is fitting. ANT is the process of tracing and understanding a complicated sociotechnical system through the interactions of its members. There are several keys to understanding ANT. First, elements of the actor-network are both actors and networks. In the words of Michel Callon, an actor-network is "simultaneously an actor whose activity is networking heterogeneous elements and a network that is able to redefine and transform what it is made of" (Callon, 1987). Second, an actor-network considers and analyzes human and non-human elements equally. ANT proponents describe the distinction between individuals, social institutions, machines, and texts as an "unnecessary duality" (Cressman, 2009). The human and non-human elements of the distance debate can only be considered as they relate to and influence one another, making ANT an appropriate framework.

ANT is not without its detractors. Some, such as Cressman, believe that frameworks based in case studies (such as ANT) minimize the social values and norms imbued in a sociotechnical system (2009). Additionally, applying a single, general framework to the vast array of uniquely constructed networks can sometimes be overly reductionist or simplistic. To address these concerns, the STS Research Paper will use the methodology of wicked problem framing. Further explained in the Method section, this methodology is tailored to sociotechnical

problems with complicated relationships and moral implications. Complementing ANT with wicked problem framing is an adept way to overcome valid critiques of ANT.

Significance of Research:

Wisely addressing the distance issue is imperative to both PGA Tour and USGA officials. Both organizations are facing increasing scrutiny for failing to setup tournament courses that can adequately challenge the best players. To quote Adam Scott, "they haven't figured out yet that long means nothing to us…I'm not challenging them to build longer golf courses; I'm challenging them to build smarter golf courses" (Marksbury 2019). As players start voicing dissatisfaction, the concern is that fans and viewers will begin to sharing in this disdain of a key element of the PGA Tour product. This discontent could have severe ramifications on television ratings, PGA Tour revenue, and eventually even recreational participation rates. For the overall health of both professional and amateur golf, the response of the PGA Tour and the USGA to distance gains must be effective. After gathering research in the Fall 2019 semester, a formal STS analysis of this topic will be conducted in the Spring 2020 semester.

Research Question:

How should the governing bodies of golf such as the USGA and the PGA Tour address distance gains on the PGA Tour?

Method:

To answer this question, the methodology of wicked problem framing will be used. Wicked problem framing first arose from a 1973 paper on urban planning by two UC Berkeley professors (Rittel and Webber, 1973). Wicked problem framing addresses societal problems that are inherently different from scientific problems and cannot be solved scientifically. These

problems are harder to define, and often have ambiguous solutions. It is difficult to determine what goal a solution should even address and when a solution has sufficiently achieved this goal. Solutions are not judged empirically like a math problem, but on the long-term ethical ramifications of implementation. There is little opportunity for trial-and-error problem solving, as even minor tweaks have lasting societal traces. Wicked problem framing is a valuable methodology for assessing systematic problems with social and technological components.

Justification:

This framework is appropriate for evaluating how governing bodies of golf should respond to distance gains. While the issue is partially scientific – scientists and engineers at equipment companies come to mind – value-based assessments are also necessary. For example, the livelihood of a player who relies on driving distance could be critically affected by efforts to dial back distance. It is also hard to define goals of addressing this issue. Should governing bodies address driving distance with the hope of creating a better fan product? Is protecting the design integrity of classical golf courses preeminent? Would dialing back distance help maintain competitive parity on the PGA Tour? Wicked problem framing is an appropriate methodology as distance gains on the PGA Tour are a complex, sociotechnical issue with ill-defined goals.

Conclusion

The technical deliverable for this prospectus addresses a need for GameForge Golf, a startup in the golf data industry. This company has struggled to manage its complicated data, and requires outside technical expertise to find a quantitative link between practice drill performance and on-course scoring. The final deliverable will include both a methodology of cleansing data and a flexible system that can recommend drills based on scoring statistics, giving GameForge a truly comprehensive approach to data-driven decision-making in the golf industry.

The STS deliverable will address a growing issue in the game of golf – distance gains on the PGA Tour. Distance gains have undoubtedly altered the PGA Tour, from the perspectives of players, fans, governing bodies, equipment companies and more. Analyzing the problem along the framework of Actor Network Theory and the methodology of Wicked Problem Framing will help readers better understand the viewpoints of and associations between these entities, in addition to what is at stake as decisions are made. The deliverable will outline the causes, both technological and human, and why addressing this issue is a necessary step for the good of the game.

References

A Review of Driving Distance—2018. (2018). Retrieved from https://www.usga.org/content/dam/usga/pdf/driving-distance/2018%20Annual%20Driving%20Distance%20Report.pdf

Average driving distance on the major tours with significant innovation milestones overlaid [graph]. (2018) Retrieved from https://www.usga.org/content/dam/usga/pdf/driving-distance/2018%20Annual%20Driving%20Distance%20Report.pdf

Callon, Michel. (1987). "Society in the Making: The Study of Technology as a Tool For Sociological Analysis." In Bijker, W., Hughes, T., & Pinch, T. (eds.) The Social Construction of Technological Systems. Cambridge: MIT Press.

Cressman, D. (2009, April). *A Brief Overview of Actor-Network Theory: Punctualization, Heterogeneous Engineering & Translation*. School of Communication, Simon Fraser University.

Dethier, D. (2018, March 13). Taylormade CEO firmly opposes rolling back golf ball: "These movements will be detrimental to the game at every level." *Golf Magazine*. Retrieved from https://www.golf.com/equipment/2018/03/13/taylormade-ceo-firmly-opposes-rolling-back-golf-ball-these-movements-will-be-detrimental-game-every

Diaz, J. (2017, September 5). The Great Divide. *Golf Digest*. Retrieved from https://www.golfdigest.com/story/golf-course-architecture-the-great-divide-jaime-diaz

Gray, W. (2018a, January 23). Davis on distance: Not "necessarily good for the game." *The Golf Channel*. Retrieved from https://www.golfchannel.com/article/golf-central-blog/davis-distance-not-necessarily-good-game

Gray, W. (2018b, February 12). R&A chief on distance: “We have probably crossed that line.” *The Golf Channel*. Retrieved from https://www.golfchannel.com/article/golf-central-blog/ra-chief-distance-we-have-probably-crossed-line

Hoggard, R. (2017, November 9). Longer courses not the answer to distance problem. *The Golf CHannel*. Retrieved from https://www.golfchannel.com/news/longer-courses-arent-answer-golfs-distance-problem

Kalland, R. (2016, July 22). Making the modern golfer: How fitness and technology have changed golf. *CBS Sports*. Retrieved from https://www.cbssports.com/golf/news/making-the-modern-golfer-how-fitness-and-technology-have-changed-golf/

Kram, Z. (2019, February 27). The 3-Point Boom is Far From Over. *The Ringer*. Retrieved from https://www.theringer.com/nba/2019/2/27/18240583/3-point-boom-nba-daryl-morey

Lewis, M. (2003). *Moneyball: The art of winning an unfair game*. New York: W.W. Norton.

Marksbury, J. (2019, August 18). Adam Scott blasts Tour tournament setups: “Long means nothing to us.” *Golf Magazine*. Retrieved from https://www.golf.com/news/2019/08/18/adam-scott-blasts-tour-tournament-setups/

McClusky, M. (2014, October 30). How the Titleist Pro V1 revolutionized golf. *Golf Digest*. Retrieved from https://www.golfdigest.com/story/how-the-titleist-pro-v1-revolu

Mell, R. (2018, February 21). Nicklaus eager to help USGA rein in golf ball distance. *The Golf Channel*. Retrieved from https://www.golfchannel.com/news/jack-nicklaus-eager-help-usga-rein-golf-ball-distance

Menta, N. (2019, October 2). PGA Tour working out driver-testing kinks to create “level playing field” | Golf Channel. Retrieved October 3, 2019, from The Golf Channel website:

https://www.golfchannel.com/news/pga-tour-working-out-driver-testing-kinks-create-level-playing-field

Miller, B., De, L. M., Horovitz, R., Pitt, B., Zaillian, S., Sorkin, A., Chervin, S., ... Sony Pictures Home Entertainment (Firm). (2012). *Moneyball*. Culver City, CA: Sony Pictures Home Entertainment.

NYT 4th Down Bot. (2014, September 4). 4th Down: When to Go for It and Why. *The New York Times*. Retrieved from https://www.nytimes.com/2014/09/05/upshot/4th-down-when-to-go-for-it-and-why.html

Rapaport, D. (2019a, February 5). Brooks Koepka: "If They Change the Ball, You Won't See Any Short Hitters on Tour." *Sports Illustrated*. Retrieved from https://www.si.com/golf/2019/02/05/brooks-koepka-golf-ball-rollback-pga-tour-distance

Rapaport, D. (2019b, February 15). The Great Distance Debate: How Ever-Increasing Driving Distance is Changing Golf. *Sports Illustrated*. Retrieved from https://www.si.com/golf/2019/02/06/golf-ball-driving-distance-rollback-debate-usga

Rittel, H.W.J., & Webber, M. (1973). Dilemmas in a General Theory of Planning. *Policy Sciences, 4*(2), 155-169. Retrieved from http://www.jstor.org/stable/4531523

Ritter, J. (2018, January 8). The Modern Swing: How and why layers go full throttle. *Golf Magazine*. Retrieved from https://www.golf.com/tour-news/2017/12/21/modern-pro-swing