PuckSmart: Visualizing Ice Hockey Games using Spatiotemporal Data and Sports Analytics

Alex Pawelczyk

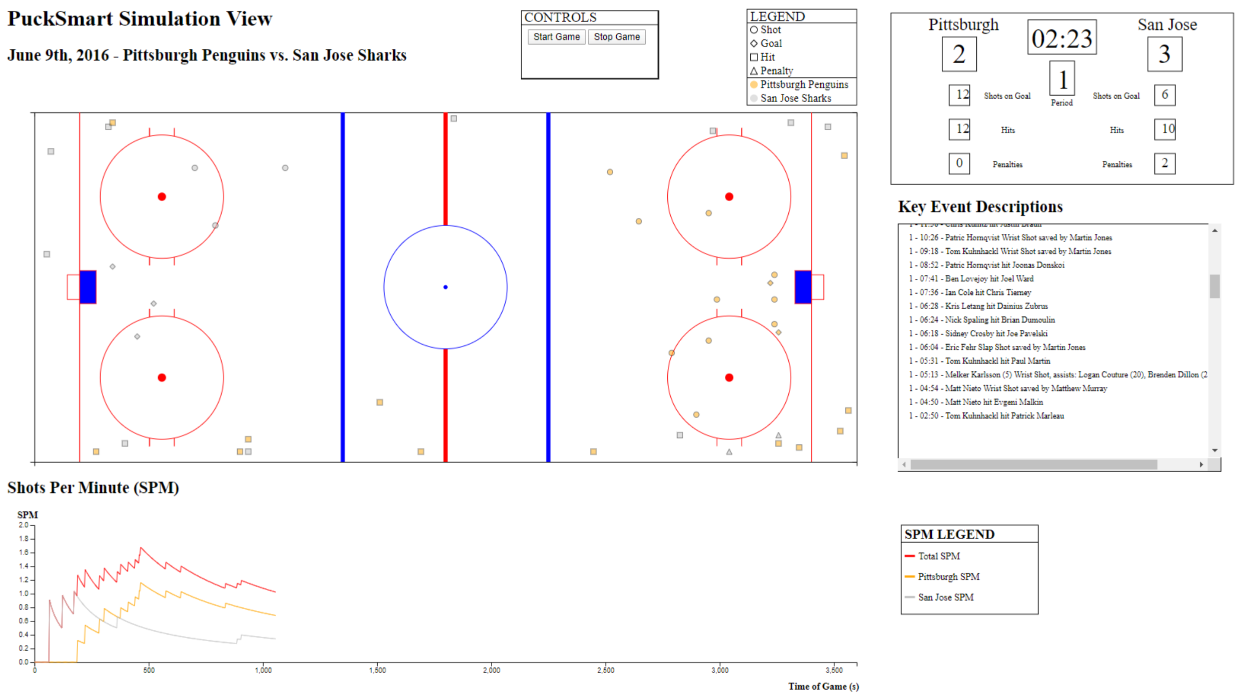

Fig. 1. Visualizing spatiotemporal data and analytics from the fifth game of the 2016 Stanley Cup Finals between the Pittsburgh Penguins and San Jose Sharks. The primary components of this visualization include a shot map, scoreboard, event description box, and shots per minute graph that dynamically update as a game is being simulated.

Abstract—A popular trend among professional sports organizations is the adoption and integration of sports analytics into daily team operations. When properly applied, sports analytics can improve the decision making process of a team, and combining these statistical methodologies with dynamic visualizations can help analysts gain comprehensive insights into games. Thus, it is important for the information visualization community to support the sports analytics industry by creating systems that enhance the existing analytical workflow. This work presents PuckSmart: a Web interface for visualizing spatiotemporal data and analytics of ice hockey games. PuckSmart converts spatiotemporal game data and analytics into meaningful visualizations of information, and the initial prototype features two different views of data. One view provides a visual summary of the key events of a hockey game, while the other view presents a novel spatiotemporal design for simulating games and dynamically visualizing their key events and analytics metrics. Domain experts were invited to analyze a sample hockey game and evaluate the effectiveness of the PuckSmart visualizations. The results of the analysis and expert feedback are summarized in two case studies.

Index Terms—Sports analytics, information visualization, spatiotemporal, PuckSmart

1 Introduction

Professional sports is a highly competitive industry in which teams strive to win championships in their respective leagues, and players aim to be the most dominant athletes at their respective positions. Wins and losses play a major role in determining the salaries and employment statuses of players, coaches, and upper-level management; thus, these individuals are always trying to find new techniques that can help them gain a competitive advantage over their opponents.

A popular trend among modern-day sports organizations involves the integration of sports analytics into daily business activities. Sports analytics refers to a collection of relevant statistics that can generate meaningful information from on-field and off-field datasets. In-game analytics focuses on measuring the performance of teams and players, while off-field analytics provides insight into business related aspects, such as increasing ticket sales or improving fan engagement. Both in-game and off-field analytics can be used to improve the decision-making process of a team, leading to better organizational performance from both sport and business perspectives.

The sports analytics market size is projected to grow from USD 1.9 billion in 2019 to USD 5.2 billion by 2024 [6], suggesting that sports analytics will play an increasingly valuable role in professional sports as technology advances. Sports analytics is valuable because it can be used to generate meaningful information related to the performance and

• *Alex Pawelczyk is with the University of Waterloo, David R. Cheriton School of Computer Science. E-mail: alex.pawelczyk@uwaterloo.ca*

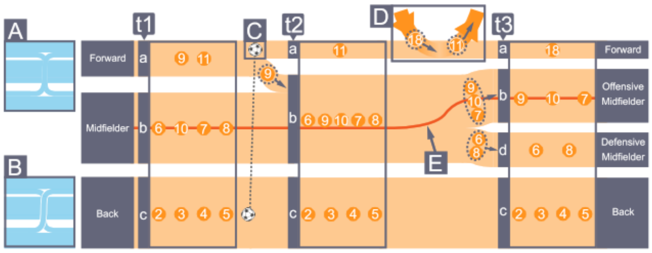

Fig. 2. Visualization of the formation changes that a team undergoes during a match. At time t1, teamB is in a 4-4-2 formation. At time t2, player number 9 drops back from a forward role to a midfielder role, resulting in a 4-5-1 team formation. At time t3, the midfielders group is divided into a defensive-midfielder subsection (player numbers 6 and 8) and an offensive-midfielder subsection (player numbers 9, 10, and 7), resulting in an overall 4-2-3-1 structure.

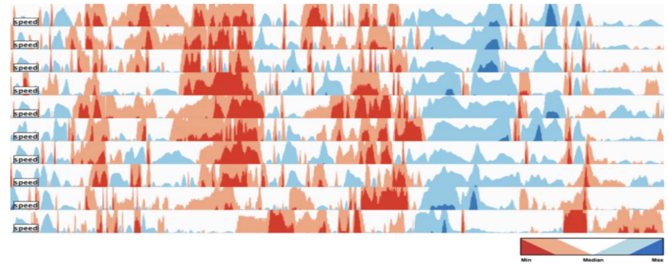

Fig. 3. Visualization of the speed of all field players of one team in the first three minutes of a soccer match.

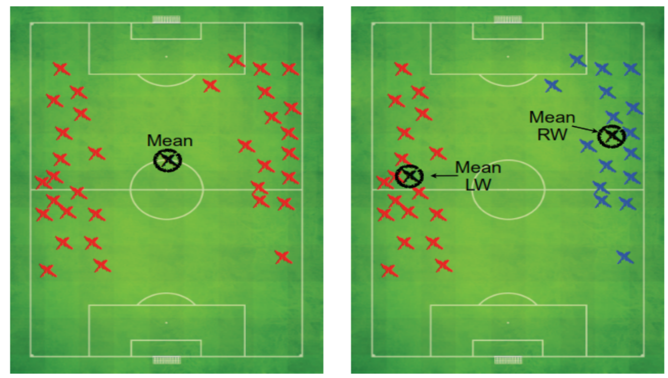

Fig. 4. Both images show the touches of a player who started a game in the left-wing position but changes halfway through a half to the right-wing. However, unlike the left image, the right image is capable of visualizing multiple roles (i.e. left-wing and right-wing) in their respective contexts.

valuation of players and teams, player fitness and safety, game tactics, fan engagement, ticket sales, and broadcast management. Ensuring that interested stakeholders can comprehend the data being presented to them is a crucial component of a sports analytics system, and information visualization techniques are effective at converting raw data into meaningful representations of information. However, information visualization is an underutilised technique in many sports analytics systems [2, 3, 18], and there exists a need for a system that combines the powers of sports analytics and interactive visualizations of game data.

The primary aim of this work is to develop a platform for visualizing ice hockey games and analytics. Specifically, this paper focuses on the visualization of National Hockey League (NHL) games and presents PuckSmart: a Web interface for visualizing NHL spatiotemporal game data and analytics. First, publicly available data for NHL games was collected via the NHL API [11]. Next, the data was cleaned to focus on the spatiotemporal attributes of key events that occur throughout a hockey game. D3.js was then used to convert the NHL spatiotemporal game data into interactive and dynamic visualizations of an ice hockey game. The initial prototype consists of a Web interface with two visualizations of data from the June 9, 2016 NHL game between the Pittsburgh Penguins and San Jose Sharks.

To evaluate the effectiveness of PuckSmart, domain experts were invited to participate in two case studies. Then, one-on-one interviews were conducted to obtain expert feedback and recommendations for improving the system. Initial feedback suggests that the visualizations used in PuckSmart can provide useful insights into the performance of individual players and teams throughout the duration of a game. However, the system is limited to the accuracy of the data it receives.

The main contributions of this work are as follows:

- An interactive visualization of spatiotemporal data that summarizes the key events of an ice hockey game and supports data exploration via filtering by different event features.

- A novel tailored design for the simulation and dynamic visualization of spatiotemporal data and analytics from ice hockey games.

2 Related Work

Recent years have seen a rise in the popularity of using visualizations to help players, coaches, upper-management, and fans better understand sporting events. This section reviews some of the related literature on the topics of sports data visualization in soccer, basketball, and ice hockey.

2.1 Soccer

Soccer analysis has been extensively studied in past years, and ForVizor [20] is an interactive visual analytic system that helps analysts conduct comprehensive investigations into spatiotemporal formation changes throughout a game. The formation view, one of the visualizations supported by the ForVizor interface, consists of a narrative timeline of the evolution of game situations (Fig. 2), a conformation matrix that shows the frequencies of formations utilized by two teams, and the distribution of formation pairs, which assists users in summarizing the difference between the formations of two teams. The interface also has a display view showing a pitch that can help with investigating player and ball movement, and a statistical dashboard that consists of several key measures for measuring team performance.

Janetzko et. al [12] present a system for analyzing high-frequency position-based soccer data at various levels of granularity, allowing for the interactive exploration and analysis of movement features and key game events. Their system deploys visualizations for single-player, multi-player, and event-based analytics, which helps the authors reach their goal of helping soccer analysts find the most important and interesting events in a match. Fig. 3 shows how the authors used horizon graphs to visualize the speed of all ten field players of one team over a three minute time period. Other features of the system include a player and ball renderer that visualizes a selected portion of a game, heat maps that can be computed for every spatiotemporal object (showing player, ball, and event positions), and line charts.

Bialkowski et. al [8] present a method that can conduct both individual player and team analysis for soccer matches. Soccer is a sport that consists of a dynamic, fast-changing environment, making it difficult to align player positions over time. The authors implement a "role-based" representation (Fig. 4) that dynamically updates each player's relative role at all frames, and then use this information to accurately visualize different strategic formations.

Similar to ForVizor, PuckSmart aims to provide a narrative timeline of the evolution of game situations. However, rather than using player tracking data, PuckSmart makes use of sports analytics and the spatiotemporal features of key game events. Moreover, the primary design goal of [12] is to help analysts find the most important and interesting

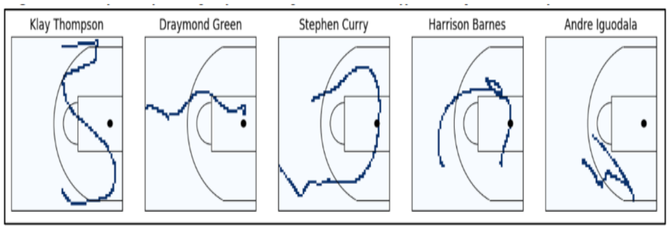

Fig. 5. Five trajectory-images for one Golden State Warriors offensive possession from the 2016 NBA Finals.

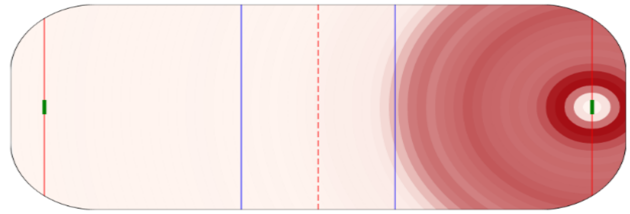

Fig. 7. Radial heat map of all shots during the 2010-2011 regular season (excluding overtime), where width of each ring equals three feet.

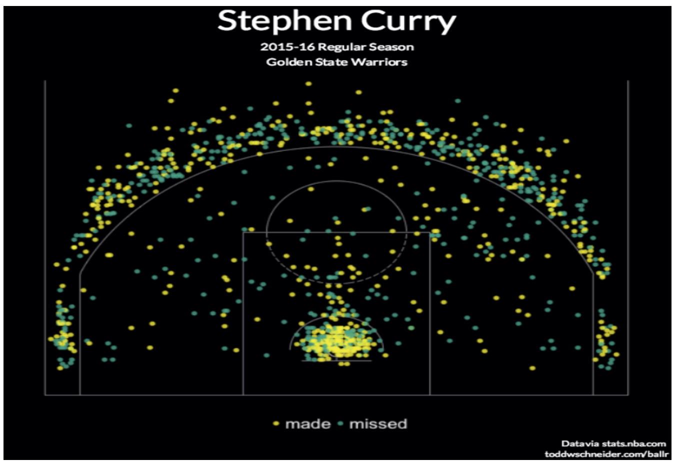

Fig. 6. Scatter chart showing all made and missed shots for one player during an NBA season. Circles identify the location on the court where a shot was taken, and color differentiates made and missed shots.

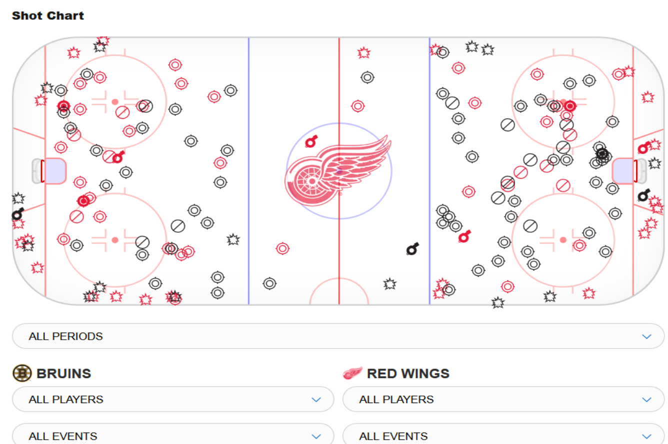

Fig. 8. ESPN Gamecenter Shot Chart for the February 9th, 2020 Detroit Redwings vs. Boston Bruins match. Different shapes represent different key events, and the placement of a shape corresponds to the location on the ice where the key event occurred.

events of a soccer game, and PuckSmart aims to do the same for ice hockey games.

2.2 Basketball

Basketball is another sport in which analytics are being rapidly adopted by professional organizations, and there have been many attempts at creating meaningful visualizations from basketball game data. For example, Nistala and Guttag [14] use a deep learning approach to generate trajectory embeddings that show how individual players move on offense. A trajectory embedding (Fig. 5) is a 32-dimensional vector of floating-point numbers that captures the semantics of a single player's movements, such as the locations of movement endpoints, screen plays, court coverage, and other spatial features.

BallR [17] is an open-source tool that obtains data via the National Basketball Association (NBA) Stats API and visualizes every shot taken by a player during an NBA season. Users are given an option to select a player and season, and based on their selections, BallR creates a customizable chart that shows shot patterns across the court. A scatter chart (Fig. 6) is one of the visualizations presented in BallR, which plots each shot as a single point of a particular color, depending on whether the shot was made or missed. BallR is an effective tool for visualizing NBA shot characteristics, and PuckSmart aims to be an effective tool for visualizing the features of NHL key game events.

2.3 Ice Hockey

Pileggi et. al [16] propose SnapShot, a visualization tool made specifically for ice hockey that displays shot data from NHL matches through a web-based user interface, primarily focusing on shot length. The system uses a dataset consisting of 72,926 total shots from the 2010-2011 NHL season to generate views such as shot maps, traditional heat maps, and radial heat maps. A radial heat map (Fig. 7) is a visualization that divides an area of ice into different sections based on how far away a shot was taken from the target net. The goal of the authors was to create a system that helps hockey analysts explore a dataspace, discuss hypotheses with other analysts, and present important findings to stakeholders.

Live covers are tools (typically a website) that display text panels and graphics containing high-level updates of key events that occur during NHL games. Live covers are used by many leagues (i.e. NHL, NBA, and MLB) and media corporations (e.g. ESPN television network, NFL network, and Fox Sports network) to present fans with an easy to understand and detailed summary of a game. Fig. 8 shows the ESPN Gamecenter Shot Chart [1], a visual summary of a NHL game where different shapes represent different events, and users have the ability to filter by event type, period in which an event occurred, and the players involved in an event.

This study primarily builds on the idea of the ESPN Gamecenter Shot Chart and proposes PuckSmart, a Web interface for visualizing the spatiotemporal features of key events in NHL games. However, unlike the ESPN Gamecenter Shot Chart, PuckSmart includes a novel simulation view (Fig. 1) that has a running clock and dynamically displays key events on the shot map as a simulation progresses. The simulation view also contains a scoreboard, a key event description box, and a shots-per-minute graph, all of which dynamically update throughout the simulation.

3 Data Processing Pipeline

This section begins with a description of the dataset that is visualized by PuckSmart. Thereafter, an outline of the data cleaning process is presented, in which the key events of a hockey game are extracted, location coordinates are converted into ones that are compatible with PuckSmart, and all necessary analytics metrics are calculated.

3.1 Data Retrieval and Description

PuckSmart depends on the collection of spatiotemporal NHL game data to visualize the key events of a game, such as shots, goals, hits, scoring chances, and penalties. This data is generated via the NHL Real Time Scoring System (RTSS), where crews of NHL statisticians watch games, make decisions in real-time, and upload data to nhl.com for the public to view [19]. Furthermore, JSON files for individual NHL games can be downloaded via the NHL API, and each file contains data about the date and time of a game, the teams and players involved in a game, player and team stats, and a sorted, play-by-play log of the key events that occurred during a game. Fig. 9 shows a sample JSON object representing the fifth game of the 2016 NHL Stanley Cup Finals between the Pittsburgh Penguins and San Jose Sharks. Most of the data needed for PuckSmart is obtained from the liveData object, which includes the sorted log of key events. The JSON file shown in Fig. 9 also represents the source of data for the initial PuckSmart prototype, where the idea is to initially focus on getting a working prototype for one game, and then gradually scale out to support more games in future prototypes.

3.2 Data Cleaning

After obtaining a dataset, the next step of the data processing pipeline involves cleaning the JSON object to get rid of any attributes that are not needed by PuckSmart (i.e., keep *liveData* and get rid of *copyright*, *metaData*, and *gameData*). The liveData object from the source dataset includes an array of 347 events that occurred throughout the game, but only 137 of these entries are key events that will be displayed in PuckSmart. Thus, the remaining 210 events were deemed unimportant and discarded.

After extracting the key game events, the next step that needs to be taken is to convert the x and y coordinates from the source dataset into ones that match the coordinates used by PuckSmart. The PuckSmart interface uses a traditional 2D computer graphics coordinate plane in which starting from zero, columns increase from left to right and rows increase from top to bottom. However, the NHL API bases their coordinate system on the dimensions of a hockey rink (200x85 feet), where the center-ice dot represents coordinates (0,0), the x-axis has a range of [-100,100] (going from left to right of the screen), and the y-axis has a range of [-42.5,42.5] (going from bottom to top of the screen). Thus, the following equations can be used to convert NHL API x and y coordinates into PuckSmart x and y coordinates, where *scale* is a variable dependant on shot map size:

$$PuckSmartX = (NHLX + 100) * scale \tag{1}$$

$$PuckSmartY = (42.5 - NHLY) * scale \tag{2}$$

Once the NHL location coordinates are compatible with PuckSmart, the final step of the data processing pipeline involves calculating the necessary statistics for the shots per minute (SPM) graph. For all seconds in a game, PuckSmart calculates the current SPM rates of the home, away, and both teams combined. At this point, the data is ready to be visually encoded into an interactive visualization.

4 Background and System Overview

The NHL is widely regarded as the best professional ice hockey league in the world, and every game generates a giant pool of raw data. This data contains information pertaining to players and important events, and from a high-level perspective, PuckSmart is a system that will convert this raw data into meaningful visualizations. This section first outlines the design goals of PuckSmart, followed by the visual encoding specifications for both the shot map and key events. Then, details are provided for the major views and interactive elements that PuckSmart displays, and the final section provides the implementation details for the first PuckSmart prototype.

4.1 Design Goals

The primary objective of PuckSmart is to visualize all key events in a NHL game. Key events include shots on goal that were saved by the opposing goalie, shots on goal that resulted in a goal, hits, and penalties.

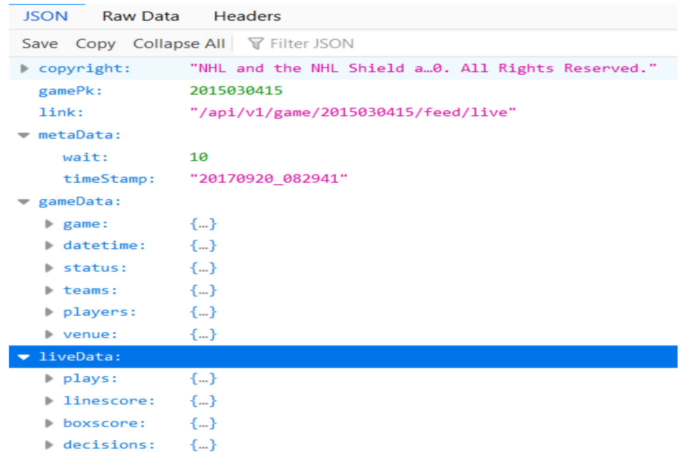

Fig. 9. General structure of a JSON object representing a NHL game [5]. PuckSmart primarily focuses on data from the liveData object, and the data cleaning process extracts all key events (e.g., shots on goal, goals, hits, and penalties) of the game.

For each of these key events, there is an associated location on ice and time of play in which that event occurred during the game. PuckSmart uses this information to produce meaningful images that show the times and locations where each event occurred on the ice surface.

A secondary objective of PuckSmart is to visualize useful sports analytics metrics that can help analysts identify interesting periods of time during a game, including situations when a team is generating a lot of shots on goal, a team does not generate a shot for a long period of time, or multiple goals are scored in a short period of time (e.g., within two minutes of each other). Knowing the periods of time when a team is either playing well or struggling can help guide coaches and players when making decisions about the amount of time that needs to be allocated towards practicing a particular skill or strategic aspect of a hockey game.

The final objective of PuckSmart is to incorporate interactive functionalities in both visualizations. Compared to static images, dynamic visualizations provide users with greater flexibility and better support for data exploration. Thus, the design of PuckSmart is centered around the use of interactive features, such as hover events, filtering, and game simulation.

4.2 Visual Encoding of Shot Map

The foundation of the PuckSmart user interface is the shot map that displays the key events of a hockey game. This map is based on the dimension and color scheme of a hockey rink defined in the NHL rule book [4], and PuckSmart converts these dimensions into pixels of the proper color to present users with a properly scaled representation of a hockey rink. To build the shot map, different kinds of scalable vector graphics (SVG) elements (e.g., lines, rectangles, and circles) are appended to the user interface via D3.js. With regards to the initial PuckSmart prototype, every 6 pixels can be thought of as one foot, meaning that the shot map is 1200 pixels wide and 510 pixels long. Future prototypes intend to generalize this map building process and implement a function that can append a shot map to a Web page at any scale. This, along with the generalized formulas for converting NHL API coordinates to PuckSmart coordinates at any scale, opens the door for additions that can improve the usability of the system, such as allowing users to customize the size of the shot map.

4.3 Visual Encoding of Key Events

To accomplish the primary design objective of PuckSmart, the output from the data processing pipeline needs to be visually encoded and appended to the shot map. A specific focus is given to the position, shape, size, and color of key event markings on the shot map, all of which are members of Bertin's original visual variables [7].

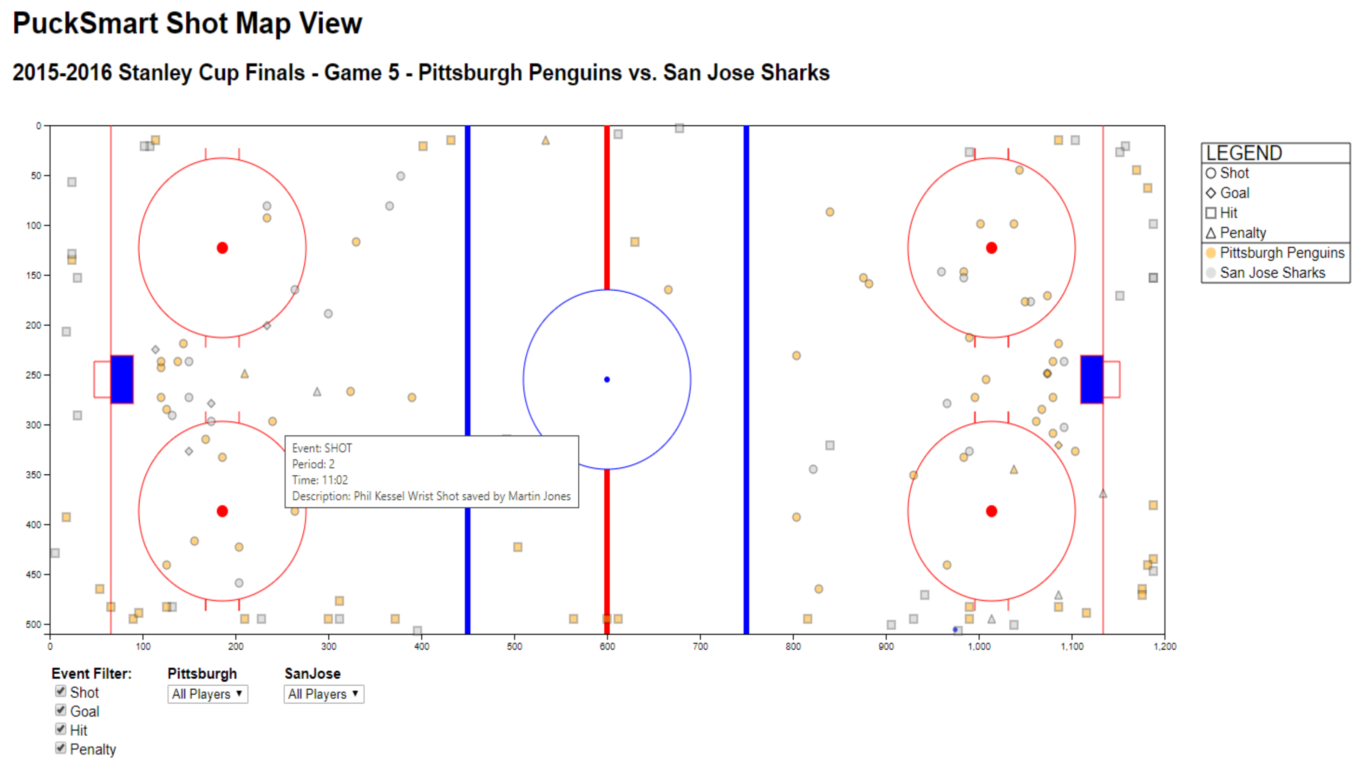

Fig. 10. PuckSmart shot map view. Key events of a hockey game (e.g., shots, goals, hits, and penalties) are displayed on the shot map, and users can filter the events by clicking on the appropriate checkboxes. A filtering by player option is also displayed, but this functionality is not yet supported by PuckSmart. The legend informs users that shapes are used to visually encode key events, and color is used to visually encode teams. Hovering over a shape displays a tooltip containing important details about an event, including the event type, period, time, and description.

Position, the most versatile and powerful visual variable [10], is a crucial aspect of both PuckSmart visualizations. PuckSmart aims to help analysts identify important spatiotemporal patterns within a game, and using the shot map as a coordinate plane for plotting key events helps visualize their locations. Positional information is useful because it can help analysts identify interesting areas of the ice to focus on, including those with high frequencies of shots, scoring chances, and goals. Using position in the PuckSmart visualizations also helps identify the distances between an event of interest compared to other events or fixed objects (e.g., how far away a shot was taken from the net).

Shape is another one of Bertin's original visual variables that is used by PuckSmart to differentiate events. Specifically, circles represent shots on goal, diamonds represent goals, squares represent hits, and triangles represent penalties. To make this symbolic meaning effective, PuckSmart presents users with a legend that defines the explicit link between different shapes and their respective meanings.

When determining the size to be used for the shapes, careful consideration was made for the tradeoff between shape size and number of overlapping events. Larger shapes are easier to see on the map, but result in a higher number of overlapping events. Conversely, smaller shapes are more difficult to see, but result in a lower number of overlapping events. Since overlapping events can make the shot map difficult to interpret and most Web browsers enable zooming in and out, PuckSmart decided to use smaller shapes. After testing different sizes, the height and width of all squares, diamonds, and triangles was set to eight pixels, along with the diameter of all circles. Also, opacity is used by all shapes to help clarify overlapping events.

An important factor in uncovering information from hockey games is knowing what events correspond to what team. PuckSmart uses color to differentiate between teams, where home-team events are assigned a default color of orange, and away-team events are assigned a default color of silver. The symbolic link between the two colors and their respective teams is defined in the legend. The main reason for choosing orange and silver as the default colors is that they contrast well amongst each other, as well as the red, blue and black colors used in the shot map. Furthermore, color was chosen to represent teams because the aim is to provide users with color blind-friendly visualizations that minimize the amount of variables that color encodes. No matter what, an NHL game will consist of two teams, meaning PuckSmart is guaranteed to only need two colors for encoding the home and away teams. However, if color was chosen to represent key events, then the current implementation of PuckSmart would require the use of four different colors, and this number would grow if future prototypes added different events to the shot map. Since most color blind people are capable of identifying different shapes, a decision was made to encode key events as shapes and teams as colors.

4.4 Major Views and Interactive Elements

The PuckSmart user interface provides users with two interactive visualizations of game data. The first is the shot map view (Fig. 10), which displays an event map with all key events for both teams. A user is able to filter these events by selecting or deselecting their corresponding check-boxes. Furthermore, hovering over an event displays a tool tip with further details, including the period of an event, the time at which an event occurred at, and a textual description of the event. The current PuckSmart prototype satisfies the requirements for the first and third design objectives, but data exploration capabilities can be increased by adding more interactive functionalities. Therefore, the plan for future PuckSmart prototypes is to add features that allow users to filter events based on specific players and times of a game, and clicking on an event will drill down into further information about an event (e.g., the stats of the player who initiated the event).

The order of key events is based on the time of game that the events occurred at, and identifying the order of key events is difficult to do through the PuckSmart shot map view. Therefore, PuckSmart also features a simulation view (Fig. 1) in which users are initially presented with a blank event map, and upon clicking the 'Start Game' button, key events begin to dynamically appear on the map whenever the key event time equals the amount of time left on the clock. Furthermore, a textual description of each event is displayed as each event occurs, providing users with a form of play-by-play commentary for the match. Unlike the shot map view, the simulation view focuses heavily on the temporal aspects of key events and dynamically displays events as they happen in real time.

To help accomplish the second design objective of PuckSmart, the simulation view also makes use of the popular SPM metric, a measure of the rate at which a team is generating shots on goal. Initially, users are presented with an empty SPM graph, where the x-axis represents the time of game (in seconds), and the y-axis represents the SPM rate. Once the simulation is started, three different colored lines begin to be appended to the graph, representing the SPM rate of the home team, away team, and both teams combined. Each point on the line graph identifies what the SPM rate was at a particular time during the game, and the graph continues to dynamically update until the end of the simulation is reached. The SPM graph is useful because it can help analysts identify important moments and trends during a game.

4.5 Implementation Details

PuckSmart is a Web application built using the HTML markup, CSS style sheet, and JavaScript programming languages. Moreover, the PuckSmart shot map and simulation views are primarily driven by D3.js, a JavaScript library that enables the binding of arbitrary data to a Document Object Model (DOM), and then applies data-driven transformations to the document to create interactive visualizations of data [9]. NHL game data is initially loaded via D3.js, and traditional JavaScript is used to identify, clean, and store the key events of interest. The final step of the data processing pipeline involves calculating the SPM statistics needed to display the SPM graph in the simulation view. Finally, D3.js is used to build interactive visualizations of NHL spatiotemporal game data, and NodeJS is used to run a simple HTTP web server and host PuckSmart locally.

5 Evaluation

This section demonstrates the evaluation of PuckSmart. Evaluations are conducted using the NHL game data presented in Fig. 9, and feedback is gathered from two hockey experts throughout the evaluation process. First, two case studies are presented in which domain experts were invited to evaluate the effectiveness of the PuckSmart shot map and simulation views, respectively. Both experts are current varsity hockey players at the University of Waterloo, each with over 15 years of experience playing competitive hockey. After the case studies, both experts were interviewed to collect their feedback and suggestions. Finally, a video analysis was conducted to determine the accuracy of the game data used by PuckSmart.

5.1 Case Study: Shot Map View

This case analyzes the visual design of the PuckSmart shot map view. Initially, Expert A looked at the shot map and legend to obtain the necessary context for understanding how key events and teams are encoded on the shot map. He then moved the mouse to the key event filters and tested the filtering functionality of the shot map by randomly selecting and deselecting the checkboxes corresponding to the four different key events. Next, the expert clicked on the Pittsburgh drop down menu, revealing a list of dummy text for player names (e.g., Player 1, Player 2, and Player 3). After clicking on the San Jose dropdown menu and seeing the same list of dummy text, the expert concluded that this feature was still in the process of being developed. Thereafter, the expert decided to filter the shot map to only display goals, and used the mouse to hover over all six goals, revealing six different tooltips that he then went on to read. The expert tried clicking on three different diamonds, and since the shot map did not change, he concluded that the shot map does not support on-click functionalities.

For further investigation, the expert clicked on the shot checkbox to add circles to the shot map. He noticed that two events occurred at the exact same location, resulting in a circle and diamond being stacked on top of each other. The expert hovered over the shapes, and a tooltip for the goal was displayed. After carefully maneuvering his mouse to the edge of the shapes, the expert was able to display the tooltip for the shot on goal. Following that, the expert decided to add hits to the shot map, and he pointed out an interesting pattern that he saw, where most hits took place outside of the faceoff dots on the outer edges of the shot map, most shots on goal were located within the faceoff circles, and all six goals were located within the faceoff circles. He explained that a typical hockey game will have most hits taking place along the boards, and goals and effective shots typically occur within the faceoff dots and a reasonably close distance away from the net (e.g., within 20 feet of the net).

5.2 Case Study: Simulation View

Expert B was invited to analyze the first ten minutes of the PuckSmart game simulation view. At first, the expert examined the layout of the user interface, and after obtaining context from the empty shot map, scoreboard, key event descriptions box, and SPM graph, he started the game simulation. The expert noticed the time on the scoreboard began to decrease in a second-by-second manner, but nothing else had changed. Then, between seconds 10 and 17 of the game, the expert identified three hits that dynamically appeared on the shot map. He also noticed the hits counter on the scoreboard had been updated to show three hits, and the key event descriptions box contained three textual descriptions of the hits. The expert read the descriptions, and then hovered his mouse over each event on the shot map to display tooltips. He concluded that the tooltips contain the same information as the key events description box. Similar to Expert A, he also tried to click on the events and came to the conclusion that PuckSmart does not support this feature.

The expert continued to monitor the simulation view for new events, and at second 64 of the simulation, he identified a San Jose goal that was appended to the shot map. He also saw the appropriate goals counter update from zero to one, and a textual description of the goal was now visible in the key event descriptions box. However, unlike the first three events, this goal caused the silver and red lines on the SPM graph to spike up, and since these lines were overlapping, the analyst concluded that the first shot on goal in the game was a San Jose goal. He also approximated the SPM rate at second 64 to be 0.91. By the end of the simulation, Expert B was able to identify an interesting area on the SPM graph where the Pittsburgh SPM line intersected with the San Jose SPM line. He correlated this intersecting area to the time of game where after facing a two goal deficit, Pittsburgh scored two goals to come back and tie the score. The expert noted that after this point, Pittsburgh appeared to take over the game, as their SPM line was trending upwards, and the San Jose SPM line was trending downwards.

5.3 Expert Feedback

After the case studies, one-on-one interviews were conducted with both experts to obtain their feedback about the design of the visualizations and any suggestions that they have for improving the PuckSmart system. Overall, the experts thought highly of PuckSmart and were satisfied with its visual design, interactive features, and use of sports analytics metrics.

Visual Design: Experts were particularly impressed by the design of the shot map due to its intuitiveness. The experts approved the encoding of key events as shapes and teams as colors, noting that these choices made it easy to identify different events for a particular team. Furthermore, both experts agreed that the legend provides all the necessary context to know that the shot map is displaying the location of all key events for both teams. Expert A recommended adding a scoreboard to the shot map view to help visualize the results of a game. He also thinks that adding a section for displaying the stats of all players would be beneficial when analyzing a game. Expert B approved of the simulation view layout, where it was easy to comprehend the purpose of all components being displayed and gather insights into the game.

System Interaction: According to expert feedback, both visualizations had useful interactive features. The experts agreed that the tooltips displayed by hovering over an event provide useful information, and they recommended adding a number to the tooltip that represents the order of a particular event. For example, the first event of a game would be labeled as 'Event 1', and the last event of a game would be labeled as 'Event *N*', where *N* equals the total number of events in a game. Expert A approved of the filtering by key event feature and liked the flexibility that it provided. However, he thinks that the current shot map view is limited in its data exploration capabilities, and recommends implementing filtering by players and times of interest. This would simplify the process of finding interesting information, such as the amount of shots on goal that player of interest had during an interesting period of time (i.e. the last five minutes of a game). Expert B was impressed with the interactive features of the simulation view, where a dynamically updating scoreboard, shot map, descriptions box, and SPM graph helped him make real-time insights into the game. However, he recommended adding filtering options to further enhance the data exploration capabilities. Expert B also recommended a feature that would allow a user to click on an event to display a window containing detailed statistics about the player who initiated the event. Finally, both experts agreed that adding animations when updating the shot map would increase the aesthetic appeal of the PuckSmart visualizations.

Sports Analytics Metrics: Both experts agree that sports analytics are an important tool for helping players and coaches gain insights into their in-game performance. Expert B had high praise for the dynamic SPM graph displayed in the simulation view, and he was able to identify important trends in the game, such as the first couple minutes where San Jose scored two goals with only three shots on goal. However, making accurate interpretations of the graph was difficult to do, and he recommended adding a functionality that displays a tooltip containing the current time of game and SPM rate for the point on the line that the mouse is hovering over. Expert A recommended adding a SPM graph to the shot map view, which currently makes no use of sports analytics metrics or visualizations. Moreover, both experts think that future prototypes should add more sports analytics metrics to the interface, such as hits and attempted shots per minute.

5.4 Video Analysis

Since the NHL RTSS has been criticized in the past for providing inaccurate spatiotemporal data [19], a brief video analysis was conducted to see if the events in the PuckSmart simulation view line up with the video feed of the June 9th, 2016 Stanley Cup Finals game between the Pittsburgh Penguins and San Jose Sharks. The video was obtained from YouTube, and as the first three minutes of the game were being played, the time, type of event, and event description for all the events was recorded. Next, the simulation view was ran for three minutes and the same information was recorded for all events. Thereafter, both lists of events were examined to identify any discrepancies between the source dataset and video feed. Table 1 summarizes these results, where only three events had all four features labeled accurately.

6 Limitations and Future Work

One of the primary limitations of PuckSmart stems from the unreliable collection of spatiotemporal attributes related to key events, where initial evaluations show that the times and locations of key events do not always coincide with the events from the video feed of an NHL game. The initial evaluations also suggest that crediting players with certain events (e.g., hits) is up to the interpretation of human analysts, implying that the accuracy of PuckSmart is dependant on the accurate collection of NHL spatiotemporal data.

PuckSmart is also limited by the size of the screen that a visualization is being depicted on. Larger screens make it easier to differentiate between the shapes and colors of key events, and there is less overlap among key events. However, a smaller screen size could result in a cluttered visualization in which many events are overlapping and it is difficult to differentiate between key events. This implies that PuckSmart may not be applicable to smartphones or other devices that have small screens. Future work will focus on conducting usability evaluations to determine the lower-bound of the screen size in which PuckSmart can depict meaningful representations of NHL game data. Further research will also be conducted towards developing an effective way to visualize overlapping key events because no matter the size of the screen, there will always be potential for overlapping events to occur (e.g., two events with the same x and y coordinates will always overlap, regardless of the size of the screen or event shape).

Table 1. NHL data accuracy for first three minutes of the evaluated game. Cells marked as true (T) indicate accurate event features, while cells marked as false (F) identify features that were not accurately recorded by NHL statisticians.

Event ID	Time	Type	Description	Location
1	F	T	T	F
2	F	T	T	T
3	F	T	F	T
4	T	T	T	T
5	F	T	T	T
6	F	T	T	T
7	F	T	T	T
8	T	T	T	F
9	T	T	T	T
10	T	T	T	T
Total Trues	**4**	**10**	**9**	**8**

Another limitation of PuckSmart comes from a lack of available technology and data for real-time player tracking. As previously mentioned, data for every NHL game is manually collected via NHL statisticians, and asking humans to manually collect player tracking data is an unfeasible request. However, the NHL plans to take advantage of recent technological advancements and deploy a system for the 2020-21 season that uses real-time puck and player tracking technologies. This system is capable of tracking the puck at a rate of 2,000 times per second in real-time with inch-level accuracy, and players can be tracked at a rate of 200 times per second [15]. Future work hopes to incorporate the data generated by the NHL tracking system into PuckSmart to create dynamic, real-time visualizations of important information. For example, a "tracking simulation" view can be developed, where rather than dynamically displaying events as they occur during a game, users are presented with a shot map that tracks the real-time position of the puck and current players that are on the ice. Other ideas include using trajectory embeddings to show player movement during a shift (similar to Fig. 5), horizon graphs for visualizing player and puck speeds (similar to Fig. 3), and heat maps for visualizing the areas of a net where goals are being scored (e.g., top-right corner of the net).

Both experts recommended adding more sports analytics metrics to future PuckSmart prototypes, and future research will focus on developing a new hockey analytics metric: *momentum*. The Oxford Dictionary of Sports Science and Medicine [13] defines psychological momentum as, "the positive or negative change in cognition, affect, physiology, and behavior caused by an event or series of events that affects either the perceptions of the competitors or, perhaps, the quality of performance and the outcome of the competition." Throughout a hockey game, there will typically be shifts in momentum between the two teams that are playing, where one team will dominate the pace of play for one period of time (i.e. have lots of shots, hits, time of possession, etc.), and the other team will dominate the pace of play during a different period of time. Currently, PuckSmart uses the SPM metric in the simulation view, and this provides good insight into the periods of time during a game that a team is getting offensive chances. The future definition of momentum will focus on combining SPM with other advanced analytics metrics to enable the real-time tracking of momentum during a hockey game.

Future work will also focus on extending PuckSmart into a team management system that can analyze and track player and team performance. The general idea is to use the visualizations and sports analytics metrics from the current PuckSmart prototype as one part of an interactive team management system. The main goal of this system is to provide a platform that can analyze, track, and visualize the performance of players and teams. Since video analysis is heavily relied upon by coaches and players, another goal is to design an interface that takes in game film as input and outputs video clips of all key events for a game. For example, this could be incorporated into the PuckSmart shot map and simulation views by displaying a video clip of on an event that a user clicks on. Other features that plan to be implemented include visualizations of individual player analytics (e.g., SPM graphs that measure the SPM rates of individual players), visualizations of team formation changes throughout a game, and a recommender system for helping coaches make lineup decisions.

7 Conclusion

This study proposes an interactive visual analytic system called PuckSmart to help analysts identify important spatiotemporal patterns and trends in ice hockey games. The PuckSmart Web interface presents users with two visualizations of NHL game data, where the shot map view summarizes the key events of a game, and the simulation view dynamically visualizes the key events and hockey analytics metrics of a game. Domain experts were invited to participate in two case studies to analyze the effectiveness of the PuckSmart visualizations, and expert interviews provide promising feedback regarding the sports analytics metrics, visual design, and interactive features of PuckSmart. The experts also see potential for a team management system that combines powerful visualizations with sports analytics, and recommend using the PuckSmart shot map and simulation views in this system.

With new NHL player tracking technologies on the imminent horizon, a surge of research activity in hockey visualization and analytics is expected. Not only will this technology help players, coaches, and upper-management gain deep insights into games, but it will also revolutionize the fan experience. For example, many hockey fans enjoy watching a game live from the arena, but struggle to understand games when watching them on a television; thus, incorporating visualizations of advanced sports analytics metrics into a television broadcast has the potential to help casual fans gain a comprehensive understanding of what is occurring during a game. Paired with a rapidly-growing sports analytics market and the global popularity of ice hockey, these new technological advancements suggest that the information visualization community should allocate future resources and interest for making meaningful contributions in hockey visualization and analytics.

References

[1] Bruins vs. red wings - game summary - february 9, 2020. `https://www.espn.com/nhl/game/_/gameId/401145896`.

[2] Sports team management software: Sap sports one. `https://www.sap.com/canada/products/sports-one.html`.

[3] Statstrack: game changing hockey intelligence. `https://statstrack.ca.`

[4] 2019-2020 nhl rule book, 2019. `https://nhl.bamcontent.com/images/assets/binary/308893668/binary-file/file.pdf.`

[5] 2020. `http://statsapi.web.nhl.com/api/v1/game/2015030415/feed/live.`

[6] Sports analytics market by sports type (individual and team), component, application (performance analysis, player fitness and safety, player and team valuation, and fan engagement), deployment model, and region - global forecast to 2024. *MarketsandMarkets*, Jan. 2020. `https://www.marketsandmarkets.com/Market-Reports/sports-analytics-market-35276513.html`.

[7] J. Bertin. Semiology of graphics - diagrams, networks, maps. 2010.

[8] A. Bialkowski, P. Lucey, P. Carr, Y. Yue, S. Sridharan, and I. Matthews. Large-scale analysis of soccer matches using spatiotemporal tracking data. In *2014 IEEE International Conference on Data Mining*, pp. 725–730, Dec 2014. doi: 10.1109/ICDM.2014.133

[9] M. Bostock. Data-driven documents. `https://d3js.org.`

[10] M. S. T. Carpendale. Considering visual variables as a basis for information visualisation. 2003.

[11] Erunion. Nhl api, Jan 2020. `https://github.com/erunion/sport-api-specifications/tree/master/nhl.`

[12] H. Janetzko, D. Sacha, M. Stein, T. Schreck, D. A. Keim, and O. Deussen. Feature-driven visual analytics of soccer data. In *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 13–22, Oct 2014. doi: 10.1109/VAST.2014.7042477

[13] K. Michael. The oxford dictionary of sports science and medicine. *Journal of Sports Science and Medicine*, 6, Mar 2007.

[14] A. Nistala and J. Guttag. Using deep learning to understand patterns of player movement in the nba. In *2019 MIT Sloan Sports Analytics Conference (SSAC)*, Mar 2019.

[15] S. Ogus. Nhl's puck and player tracking technology tested in all-star game; set to be up and running for playoffs. *Forbes*, Jan 2020. [Online]. Available: `https://www.forbes.com.` [Accessed Mar. 4, 2020].

[16] H. Pileggi, C. D. Stolper, J. M. Boyle, and J. T. Stasko. Snapshot: Visualization to propel ice hockey analytics. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2819–2828, Dec. 2012. doi: 10.1109/TVCG.2012.263

[17] T. Schneider. Ballr: Interactive nba shot charts with r and shiny, Mar 2016. [Online]. Available: `https://toddwschneider.com/posts/ballr-interactive-nba-shot-charts-with-r-and-shiny/` [Accessed Apr. 6, 2020].

[18] H. K. Stensland, V. R. Gaddam, M. Tennøe, E. Helgedagsrud, M. Næss, H. K. Alstad, A. Mortensen, R. Langseth, S. Ljødal, O. Landsverk, C. Griwodz, P. Halvorsen, M. Stenhaug, and D. Johansen. Bagadus: An integrated real-time system for soccer analytics. *ACM Trans. Multimedia Comput. Commun. Appl.*, 10(1s), Jan. 2014. doi: 10.1145/2541011

[19] T. B. Swartz. Hockey analytics. *Wiley StatsRef: Statistics Reference Online*, p. 1–10, 2017. doi: 10.1002/9781118445112.stat07965

[20] Y. Wu, X. Xie, J. Wang, D. Deng, H. Liang, H. Zhang, S. Cheng, and W. Chen. Forvizor: Visualizing spatio-temporal team formations in soccer. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):65–75, Jan. 2019. doi: 10.1109/TVCG.2018.2865041