Adaptive Agent for Player-Specific Fitness and Health Incentives in Mobile Location Based Games

Spencer Frazier, Chao Huang, Sarit Kraus, Yu-Han Chang and Rajiv Maheswaran

University of Southern California
4676 Admiralty Way #1001
Marina Del Rey, CA 90292

Preface and Relation to Data-Driven Wellness

As location-based mobile games become more popular, movement becomes an integral part of game play. This provides an opportunity for the game to influence player behavior in the real world, potentially inducing more physical activity (and better health) through intelligent adaptation of the game mechanic. We describe the application of Markov Decision Processes (MDPs) to model the player's behavior in a custom-built location-based zombie fighting game. The game agent uses this model - a user specific optimal policy (USOP) - to adjust the game behavior to encourage as much game play as possible. Our experiments with human subjects showed that game play time was indeed increased over the control condition. We look at how games can be used to model user behavior and then unobtrusively effect agent-determined behavioral change. We believe that to gather meaningful data and to maintain user engagement, a game must not only be robust in the types of data gathered but also be well-developed and well-designed. Researchers have recently begun to explore using mobile games to incentivize healthy living habits (Buttussi and Chittaro 2010), through active playing (Verma 2005) and passive encouragement (Pollak et al. 2010). In this paper, we use past play data to model the user and adapt the game's actions to produce more user activity. This kind of personalization in games has been studied before, using procedural content generation (Shaker 2010), genetic algorithms (Verma 2005), Bayesian knowledge tracing in educational games (Molnar 2011), and affective computation that models emotions (Hudlicka 2011). Here we parametrize the spatio-temporal action space of the game agent to enable it to adjust its actions based on a learned policy, treating the optimization problem as an MDP.

The Game Model

The game, ZOMBIESC , is a location-based mobile game where users must move close to objects in order to interact with them; in this case, the range was 50m. There are 3 types of objects: small, medium, and large zombies. The small zombies give a reward of cash and tickets simply by tapping on them once close enough. Tapping the medium and

large zombie icons initiates a battle screen, in which users use weapons (tomatoes, baseballs, or trash) to defeat multiple advancing zombies. Again, success is rewarded with cash for items and tickets for a raffle. To adapt to the player, the game has a parametrized action space that describes the placement of the goals (zombies) around a player's current location. The first variable describes the average distance of zombies from the user, and has three possible values: within range, just out of range, and far away. The second variables describes content clustering. In the unclustered setting, zombies are placed uniformly on a disk of radius 500m around the player. In the clustered setting, zombies are placed mainly within three smaller 200m disks within the larger 500m disk. The third variable adjusts the ratio of small vs medium and large zombie objects. Finally, the fourth variable determines whether or not nearby objects are deliberately skewed away from the player's current heading, i.e. when this variable is true, a player would be forced to deviate from their current path in order to interact.

The Player Model

The game currently models the players as a single group, using a small five-state MDP. The states are parametrized by two variables, denoting movement (moving (M) and not moving ($\neg M$)), and state of play (actively playing (A), playing (P), and not playing ($\neg P$)). The game infers the value of these variables by using observable sensor readings during each session of play: total time played, number of small zombies engaged, number of large zombies engaged, distance travelled based on regular GPS updates, and number significant shifts in bearing. If a user has not interacted with any content, that player can be considered as not playing. If a user has interacted with any content at all but has not moved, then that player can be considered to be in the ($P\neg M$) state. If a user has interacted with content and traveled a distance greater than 60m (equivalent to 3 significant location up-

Scenario and Value	Hours State 1	Dist. (km) State 1	Hours State 2	Dist. State 2	Hours State 3	Dist. State 3
Wk 2 USOP μ	2.80	0.18	0.37	2.19	0.07	0
Wk 2 USOP σ	4.53	0.30	0.58	5.89	0.21	0
Wk 2 USOP Z	3.25	0.07	0.03	7.85	0.02	0
Wk 2 USOP P	0.99	0.52	0.51	1	0.51	0.5
Wk 2 Control μ	3.72	0.33	1.67	14.51	1.98	6.12
Wk 2 Control σ	7.19	0.23	2.82	22.20	3.8	13.73
Wk 2 Control Z	10.59	0.06	1.44	7.84	2.68	3.00
Wk 2 Control P	1	0.53	0.93	1	0.99	0.99

AVERAGE DISTANCE(KM) EACH PHASE

AVERAGE TIME (HOURS) EACH PHASE

dates) we consider them to be in the (PM) state. We set this distance to 60m to account for GPS noise. A player is said to be actively playing is they have interacted with more than 10 small content, more than 3 large content, travelled more than 60m, undertaken 5 bearing shifts greater than 45 degrees and played for 10 minutes or more.

The game agent acts by choosing a new distribution of goals (zombies) every 60 seconds. It first gathers data from players interacting with randomly generated actions, and uses this data to estimate the transition matrix, i.e. the values $p(s, a, s')$, where s is the current state of the user, a is the action chosen by the game agent, and s' is the next state of the user. We assign positive rewards to the states where the user is moving and playing, and even larger rewards to the actively playing state. We can then solve for an optimal policy.

Experiments

We recruited 40 players to play the game for 2 weeks. The players were incentivized by a raffle. The game client had no learning agent during the first week. During the second week, players were randomly distributed into a control and treatment group, controlling for total play time during week 1. The control group received the same game client as the week before, and the treatment group received a game that alters its behavior in response to the user state, based on the policy learned from week 1 data.

During week 1, 38 players engaged in 873 battles and interacted with 3800 zombie objects on the map. In total the players accumulated 88km of travel over 20 hours of total play. During the second week, players were evenly divided into two groups, but only 17 players actually continued with the experiment. Of these, 11 players were in the treatment group, and 6 players were in the control group. The treatment group totaled over 1000 battles, 3300 map object interactions and 27 km traveled. The control group totaled 200 battles, 900 map object interactions, and 25km traveled. In Table 1 we demonstrate a significant shift in play occurred after treatment application but only in the first state. We understand Z to measure standard deviation and P to denote probability where bother relate to the standard normal distribution.The amount of total play in state 1 increased while the total distance traveled in this state decreases. This is valid because the state 1 corresponds to our NMP - not moving, still playing - state. Though the distance travelled and time spent in the other states in comparative ratios of play did increase in states 2 and 3 - these increases were not statistically significant. The p-score for state 1 one demonstrates a statistically significant gain in total play hours spent in that
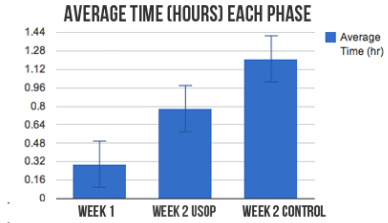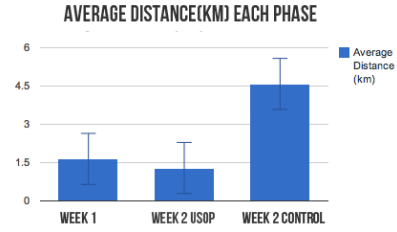
state. The distance exhibits an even higher p for this same reason.

In Figure 2 we show the total distances traveled between week 1, week 2 control and week 2 treatment with their error ranges. In Figure 3 we show the total time spent playing the game in week 1, week 2 control and week 2 treatment.

Conclusion

We were unsuccessful in motivating players to spend more time in the actively playing (A) state though we were more successful in motivating them to spend more time in the $\neg MP$ and MP states. It seems that by adapting to a user's preference in terms of content distribution, we may have inadvertently incentivized users to remain stationary where the random distribution has some feature that better incentivizes movement. Closer examination of the data will hopefully allow us to extract a more refined action space so that we can provide map-based goals that incentivize players to transition from non-moving to moving states.

References

Buttussi, F., and Chittaro, L. 2010. Smarter phones for healthier lifestyles: An adaptive fitness game. In *Pervasive Computing*.

Hudlicka, E. 2011. Affective computing for game design. In *4th Intl. North American Conference on Intelligent Games and Simulation*.

Molnar, A. 2011. Educamovil: Mobile educational games made easy. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications*.

Pollak, J.; Gay, G.; Byrne, S.; Wagner, E.; Retelny, D.; and Humphreys, L. 2010. Time to eat - using mobile games to promote healthy eating. *IEEE Pervasive Computing* 9:21–27.

Shaker, N. 2010. Towards automatic personalized content generation for platform games. In *Proceedings of the Sixth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*.

Verma, M. 2005. Personalised gaming: a motivation and overview of literature. In *Proceedings of The 8th Australasian Conference on Interactive Entertainment: Playing the System*.