EV-Human: Human Localization via Visual Estimation of Body Electronic Interference

Xinfeng Li*, Jin Teng*, Qiang Zhai*, Junda Zhu†, Dong Xuan*, Yuan F. Zheng† and Wei Zhao‡

*Dept. of Computer Science and Engineering, The Ohio State University, USA. {lixinf, tengj, zhaiq, xuan}@cse.ohio-state.edu
†Dept. of Electrical and Computer Engineering, The Ohio State University, USA. {zhuj, zheng}@ece.osu.edu
‡The University of Macau. weizhao@umac.mo

Abstract—**Human localization is an enabling technology for many mobile applications. As more and more people carry mobile phones with them, we can now localize a person by localizing his mobile phone. However, it is observed that presence of human bodies introduces heavy interference to mobile phone signals. This has been one of the major causes of inaccurate wireless localization for humans. In this paper, we propose using video cameras to help estimate human body's interference on mobile device's signals. We combine human orientation detection and human/phone/AP relative position inference estimation to better measure how a human blocks or reflects wireless signals. We have also developed a signal distortion compensation model. Based on these technologies, we have implemented a human localization system called EV-Human. Real world experiments show that our EV-system can accurately and robustly localize humans.**

I. INTRODUCTION

Human localization is an important technology for a variety of applications such as location-based services [6]. The growing popularity of mobile phones provides us an opportunity to localize a person by localizing the mobile phone he/she is carrying. A large category of traditional wireless localization is infrastructure-based localization. Wireless infrastructure such as Base Stations (BSs) or Access Points (APs) collects wireless signals sent by the mobile phone and performs triangulation in order to identify a person's location commonly based on the Received Signal Strength Indications (RSSIs) [1].

However, localization of a mobile phone that is on a human body can be very inaccurate. One of the most important causes is severe human body interference on wireless signals sent by mobile phones [13]. For example, when the human body totally blocks wireless signals, the RSSI reading at the receiving AP end will undergo a decrease of up to 15–20 dB [13]. This interference severely degrades the accuracy of the traditional RSSI-to-distance conversion based on the free space attenuation model and leads to huge localization errors. Conventional wireless localization techniques cannot automatically detect and rectify this RSSI distortion unless certain information is reported by cooperative people, e.g., their orientations and body movements, which was studied in SpinLoc [8].
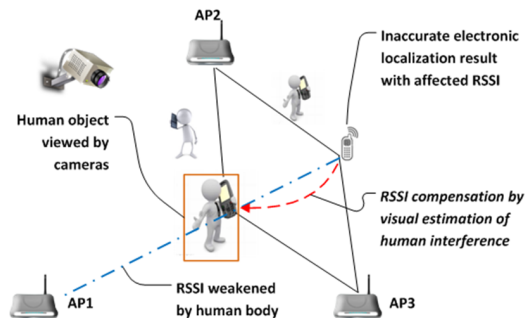
Fig. 1: Video cameras help estimation of human body interference on electronic signals.

In this paper, we propose a methodology of utilizing video cameras to automatically detect and compensate for this signal distortion due to human body interference for accurate human localization. As shown in Fig. 1, we want to localize people with infrastructure including video cameras and APs. Our basic idea is to "see" people's position and posture via video cameras and obtain clues on how the body blocks or reflects wireless signals to cause interference and measurement errors at the AP side. Human body interference on electronic signals is largely determined by the relative position between the mobile device, the human body, and the AP [13]. It is almost impossible to reliably know this relative position without actually "seeing" the scenario. So we seek help from video cameras to detect human orientation with respect to the AP and the phone's relative position on the human body. Thus human body interference can be estimated accordingly and RSSI measurements can be amended according to a proper compensation model. This can achieve much more accurate localization of humans.

As simple as the idea may sound, there are two challenges that we must face. The first challenge is finding the relative positions between the mobile phone, the human body, and the AP, which enables finer estimation of human interference. To achieve this, direct sight of phones is best in order to estimate their relative positions. However, this is not always possible as people may place their phones in their pocket or bag. The second challenge is rectifying the wireless signal distortion due to human interference. That is, how much should we compensate to a distorted RSSI, given the known or inferred human-phone position? Although human interference on electronic signals

has been observed by existing works [8], [13], no clear RSSI compensation model has been proposed.

In this work, we propose and implement a localization system called EV-Human in which we apply the above methodology of visual estimation of human body interference to our existing work, EVLoc [11]. EVLoc is a localization system that associates electronic localization results with visual localization results in order to achieve better localization accuracy. However, EVLoc is designed to localize mobile devices without considering interference from their bearers so it cannot be directly applied to localize humans. After combining the above methodology with EVLoc, the new EV-Human system can localize humans with high accuracy. We claim the following contributions:

- We have proposed the methodology of using visual signals to help estimate the human body interference on electronic signals. The estimation can help us rectify signal distortions due to human body interference, thereby greatly improving electronic localization accuracy.
- We have proposed a set of effective and practical technologies to solve the above two challenges brought by visual interference estimation. We leverage human orientations and movements to effectively infer relative positions between humans, phones, and APs. We have also derived an empirical signal compensation model given the relative position.
- We have conducted extensive indoor and outdoor experiments with static or mobile humans. The results show that our EV-system can accurately and robustly localize humans. Compared to the existing work, EVLoc, EV-Human can significantly improve the localization accuracy by $50\% \sim 100\%$.

The rest of this paper is organized as follows. Section II discusses related work. Section III introduces our methodology. Section IV details the human localization designs. Section V presents our evaluation results. Finally, Section VI concludes this paper.

II. RELATED WORK

Our work aims to perform human localization with mobile device which has not been studied extensively. Many localization techniques [3], [5] implicitly ignore the impact of human bodies, assuming the electronic devices or sensors are isolated. Scene analysis based localizations, for example RADAR [2], fingerprint each location of a scene for later matching with observed RSSIs. We argue that the varying orientations of human beings may change the human body interference significantly, and thus affect the accuracy of fingerprints. Device-Free Localization (DFL) [10], [12] utilizes the impact of human bodies on the electronic signals to locate the human objects. These signal-based techniques can detect the "existence" of human objects but fail to localize the held phones. The identity of the person is also of interest to us. Borealis [13] and SpinLoc [8] are two systems utilize the shadowing effect of human bodies for localizing of APs or

users themselves. However, both of them require cooperative users who have to spin. This is not feasible when the people are localized passively or unwillingly.

III. METHODOLOGY

In this section, we are going to discuss the methodology of EV-Human, i.e., how video cameras can help estimate human interference and thereby greatly enhance localization accuracy.
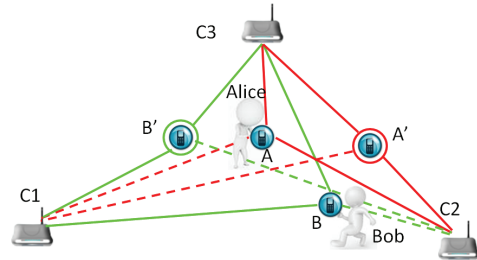

Fig. 2: Localization errors due to human interference

In the example in Fig. 2, we can see that Alice's phone is blocked by her own body in the perspective of AP C_1. So C_1 may see a weaker signal from Alice's phone than if Alice is not in the way. It translates to a longer distance between Alice's phone and C_1, i.e., from AC_1 to $A'C_1$. As a result, Alice's phone may be located at spot A' instead of A. Similarly, Bob's phone may be mistakenly localized at spot B', whereas B is the correct position. Now a phone misplacement will occur due to the erroneous localization results. We will think Alice's phone belongs Bob because it is geographically closer to Bob, while Bob's phone belongs to Alice.

In this case, the error comes from distorted RSSI readings. Without video cameras, we have no way to figure out the blocking and misplacement. But if we can see how Alice and Bob block their own phones, we may be able to correct the RSSI readings. For example, if we see that Alice stands in between her phone and an AP, we should increase the RSSI reading at that AP properly. The compensated RSSI will better reflect the accurate distance between her phone and the AP. As a result, we will be able to "pull" the localization result of Alice's phone back from A' to A, and correctly place the phone on Alice, instead of on Bob.

IV. SYSTEM DESIGN

In this section, we present the system design of EV-Human. EV-Human builds upon our existing work called EV-Loc [11] and integrates the above methodology of visual interference estimation.

A. System Workflow

Our localization process is mainly composed of three phases: 1) visual and electronic signal collection and pre-processing; 2) visual signal aided human interference estimation and signal compensation; 3) E-V object association.

The first phase is to collect visual signals and electronic signals. In this work, visual signals are the videos captured from multiple deployed cameras and electronic signals are the
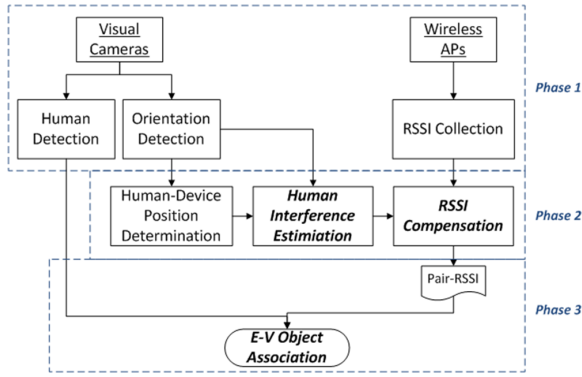
Fig. 3: System Workflow

wireless devices' WiFi RSSIs observed by several APs laid out around the area of interest. Human detection is performed to find human object locations and human orientation detection techniques are applied to obtain human orientations.

The second phase is to estimate human body interference based on visual signals. As illustrated in Fig. 4, we consider the body shadowing and reflection on electronic signals. The human body interference is related to the included angle made by the wireless device, human body and AP. We need two components to estimate the human body interference: human orientation detection and human-device relative position determination. Computer vision techniques are employed to detect the human orientation and a brute-force search with heuristic boosts is applied to find the most possible human-device position.

After knowing the relative positions among the wireless device, human body and AP, we compensate the electronic signal distortion. We derive the empirical compensation model from human interference experiments. Note that we only know the interference caused by each person but do not know which device is actually affected by that person. Thereby, we need to associate electronic (E) objects with visual (V) objects, i.e., which person owns which phone, in the third phase.

To find the correct association, we leverage the fact that the electronic localization result of a phone and the visual localization result of the owner is the shortest. Thus the Hungarian algorithm [7] is used to find the best association with the minimum distance sum of all E-V pairs. Different from EVLoc, we feed the RSSI difference between an AP pair, called "Pair-RSSI", rather than their absolute RSSI values to the Hungarian algorithm. The intuition behind Pair-RSSI is the observation that the RSSI difference is fairly obvious and stable, especially when a person is facing one of the APs. The Pair-RSSI captures the human impact robustly and is beneficial for canceling the impact of phone heterogeneity, environmental interferences and system biases.

B. Visual Signal Aided Human Interference Estimation

In this section, we describe how to perform human orientation detection and human-device relative position determination. Then we discuss our empirical electronic signal compensation model.
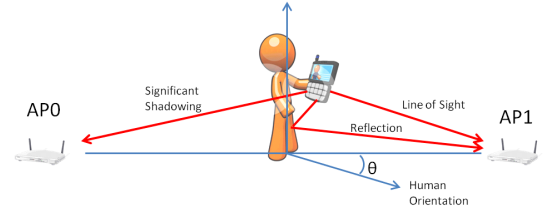

Fig. 4: Human body interference illustration

1) Human Orientation Detection: We employ two computer vision techniques to detect human orientation: human tracking and face detection.

-Human Tracking: For the general cases in which the persons are continuously moving, we use visual human tracking techniques to track the movement of all human objects and obtain their facing orientations consecutively. Considering that people usually walk in the forward direction, we can determine their orientations based on their moving directions.

We employ multiple cameras from different angles and fuse the information from all the cameras' views to keep continuous track of all human objects using the K-Shortest Path (KSP) tracking algorithm [4]. Further, we employ third-order spline interpolation to smooth the output from visual tracking such that the resulted trajectories are continuously differentiable. The persons' orientations are obtained by calculating the tangential directions of the smoothed trajectories at the sampled locations.

-Face Detection: When a human object is static or stays still, the human face can be an alternative for determining human orientation. The Viola-Jones face detection algorithm [9] is performed on every captured images. When a person is facing the camera with a small deviation angle, his frontal face will be able to be detected and thus his orientation is determined. As face detection requires strict frontal view with a bias angle $\leq 15°$, multiple cameras are needed in order to cover the $360°$ of the area of interest.

2) Human-Device Relative Position Determination: Next, we need the human-device relative position to determine the included angle of the wireless device, human body and AP. An intuitive idea is to do brute-force search. By assuming one human-device relative position for each person, we can compensate the signal distortion and perform E-V object association. After examining every case, the best matching result with the lowest "distance" is selected as the final localization result. However this approach is highly time consuming. Here we provide two ways to boost it.

First, the human orientation detection result is utilized to provide a reference direction. We assume the human-device relative position keeps the same during the localization period. Thus, the included angle of the wireless device, human body and AP is uniquely determined by the human orientation. As people usually put the phone in front of body, in left or right pocket or in the knapsack at the back, we consider these four relative positions in our system. Second, we let Pair-RSSI play as a filter during the search. The idea is to check the validity of "Pair-RSSI" values. For example, the "back" position can
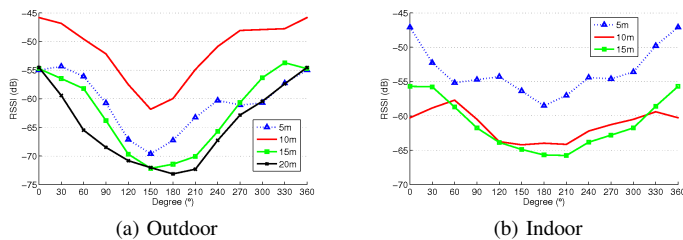
(a) Outdoor   (b) Indoor

Fig. 5: Human Interference against Angles

be safely excluded if all RSSIs observed by an AP in the front are much higher (for example, 10 dB) than those observed by another AP in the back.

3) Electronic Signals Compensation: We conducted experiments to measure the human body interferences on electronic signals. Two APs are placed on a line, as is shown in Fig. 4. One person holds a phone in front of his chest and stands at an angle of θ to the connecting line. We measure the RSSI readings on the two APs when θ varies. The results are shown in Fig. 5. We observe that when the person faces one of the APs, the AP behind the person receives a significantly weaker signal, e.g., 15-20 dB. Meanwhile, the wireless signals from the mobile phone bounce off the human body and the reflected signal reach the AP in the front. In our experiments, the AP in front of the person *usually* receives a slightly stronger signal, e.g., 3-5 dB.

We have modeled and calculated the RSSI differences at AP0 and at AP1 as follows (refer to Appendix for details).

$$\Delta RSSI_0 = \alpha \log \frac{akc + b\sqrt{a^2k^2 - c^2 + b^2}}{akc - b\sqrt{a^2k^2 - c^2 + b^2}}$$

$$\Delta RSSI_1 = \beta \log \frac{\sqrt{(l\sin(\theta))^2 + (c + l\cos(\theta))^2}}{l}$$

TABLE I: Standard deviation of RSSI vs. Pari-RSSI (dB)

	AP1	AP2	AP3	AP1-AP2	AP2-AP3	AP3-AP1
σ	2.40	3.54	3.21	1.79	1.27	1.98

RSSI readings suffer from severe noise influence. Directly using RSSI as the localization measurement is not reliable. In this work, we propose using Pair-RSSI, the RSSI difference between an AP pair, instead of absolute RSSI readings. We can observe that the value of RSSI difference is outstanding and fairly stable when θ is near $0°$ or $180°$. Furthermore, Pair-RSSI is beneficial for canceling environmental noises. We show in Table I the variances of absolute RSSI and Pair-RSSI observed in our experiments.

V. EVALUATION

We have implemented EV-Human system and conducted real world experiments in human localization. In this section, we report the system evaluation results.

A. Experiment Settings

-Scenarios: We evaluated EV-Human in static mobile scenarios. In the static scenario, people keep static in one frame and change their locations and orientations in the next frame. We conducted outdoor (10.7 m × 18.3 m) and indoor experiments (5 m × 10 m) with 6 participators. We recorded 9 outdoor frames and 16 indoor frames. In the mobile scenario, the people are moving freely in the experiment scene. We conducted the mobile experiments indoors (8 m × 11 m) with 3 participators. The phones were either held in the hands or put in the trouser pockets. For each case, we recorded E and V data for about 2 minutes (400∼500 frames).

-Evaluated Algorithms: The following five algorithms are evaluated for comparison. The first one is the baseline E localization algorithm (P1) through trilateration. The second algorithm EVLoc (P2) integrates the electronic localization and visual localization. We add the human interference in the third approach (P3). Next, we incorporate the Pair-RSSI idea in the fourth algorithm (P4). Finally in the fifth algorithm (P5) we infer the human-device relative positions assuming the human-device relative positions are unknown.

-Metrics: 1) Average error distance. It refers to the average distance from the actual location to the localized position over all human objects. 2) Convergence speed. It refers to how many frames an algorithm needs before achieving correct localization results.

B. Experiment Results

1) Static Scenario: In the static scenario, we evaluated P1∼P4 algorithms given that the human-device relative positions are known (in front of the human bodies). The results are shown in Fig. 6a, 6b. We can observe that our proposed algorithm P4 has the lowest average error distance among the four algorithms and reaches zero error performance fastest. Specifically, with the help of human interference model, we can amend the RSSI deviation and reduce the average error distance by around 1 meter in one frame. Although P2∼P4 all get the correct localization results with the 9 outdoor frames' data, P2 needs all the 9 frames while our algorithm P4 only needs 5 frames. The convergence speed has been improved by nearly 100%. For the indoor experiment, neither of P2 and P3 could reach the zero error with all 16 frames' data but P4 has already found the correct localization results with 9 frames.

2) Mobile Scenario: The performances with mobile human objects are shown in Fig. 6c, 6d. We can observe similar performances to those in the static scenario. P2 performs much better than the traditional electronic localization and a little bit worse than P3 as the human interference is neglected. Furthermore, P4 performs best among all algorithms which costs 39 or 56 frames to find the 100% correct matches of EV objects. It means P2 may fail to find the correct electronic identities of the human objects in a 20 seconds (100f / 5fps = 20s) long video but our algorithm P4 can successfully pick out one person's corresponding electronic identity as long as he once walked into the monitored area for about 8∼11 seconds.

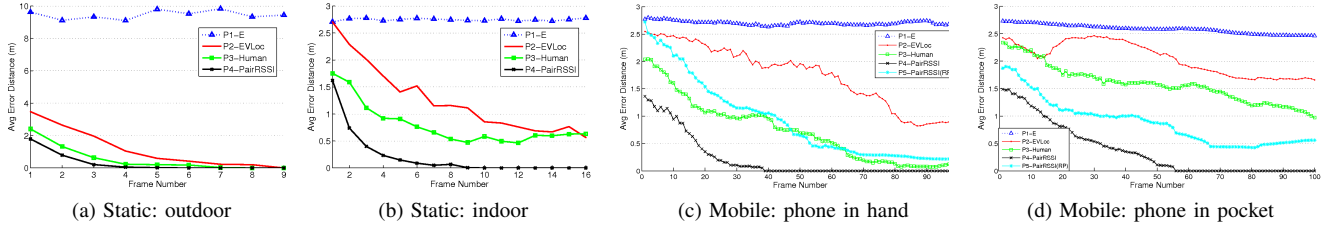| (a) Static: outdoor | (b) Static: indoor | (c) Mobile: phone in hand | (d) Mobile: phone in pocket |

Fig. 6: Performance with static and mobile human objects

Besides, we observe that P5 has a very decent performance compared with other algorithms. Although it needs 90∼100 frames to reach the accuracy of 0.2 meter, it performs much better than P2. For example, the average error distance has been decreased from 0.90 meter to 0.21 meter at the 100th frame (Fig. 6c). We also notice that the average error angle of the determined relative positions is only 57.3° or 41.0° when phones are in the hands or in the pockets. In other words, we may regard a phone locates at the left diagonal of a human body while it is actually in the front. Such an accuracy is acceptable in human localization.

VI. CONCLUSION

In this paper, we have explored the problem of wireless human localization problem, which remains a difficult problem due to human body interference. We have introduced video cameras to estimate body interference on wireless signals. An effective and practical approach to compensate the interference has been proposed. We have implemented the system called EV-Human and conducted extensive real world experiments to confirm the efficiency and robustness of our localization scheme.

REFERENCES

[1] T. A. Alhmiedat and S.-H. Yang. A survey: localization and tracking mobile targets through wireless sensors network. In *PGNet 2007, Eighth Annual Network Symposium*, June 2007.
[2] P. Bahl and V. Padmanabhan. RADAR: an in-building rf-based user location and tracking system. In *Proceedings of IEEE INFOCOM*, volume 2, pages 775 –784 vol.2, 2000.
[3] N. Banerjee, S. Agarwal, P. Bahl, R. Chandra, A. Wolman, and M. D. Corner. Virtual compass: Relative positioning to sense mobile social interactions. In *Pervasive '10*, pages 1–21, 2010.
[4] J. Berclaz, F. Fleuret, E. Türetken, and P. Fua. Multiple object tracking using k-shortest paths optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011.
[5] M. Ding, F. Liu, A. Thaeler, D. Chen, and X. Cheng. Fault-tolerant target localization in sensor networks. *EURASIP Journal on Wireless Communications and Networking*, 2007(1):19–19, 2007.
[6] I. A. Junglas and R. T. Watson. Location-based services. *Commun. ACM*, 51(3):65–69, Mar. 2008.
[7] H. W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2:83–97, 1955.
[8] S. Sen, R. R. Choudhury, and S. Nelakuditi. Spinloc: Spin once to know your location. In *Proceedings of HotMobile*, pages 12:1–12:6, 2012.
[9] P. Viola, M. J. Jones, and D. Snow. Detecting pedestrians using patterns of motion and appearance. *International Journal of Computer Vision*, 63:153–161, 2005.
[10] M. Youssef, M. Mah, and A. Agrawala. Challenges: device-free passive localization for wireless environments. In *Proceedings of ACM MobiCom*, pages 222–229, 2007.
[11] B. Zhang, J. Teng, J. Zhu, X. Li, D. Xuan, and Y. F. Zheng. EVLoc: Integrating electronic and visual signals for accurate localization. In *Proc. of ACM MobiHoc*, 2012.
[12] D. Zhang and L. Ni. Dynamic clustering for tracking multiple transceiver-free objects. In *Proceedings of IEEE PerCom 2009*, march 2009.
[13] Z. Zhang, X. Zhou, W. Zhang, Y. Zhang, G. Wang, B. Y. Zhao, and H. Zheng. I am the antenna: accurate outdoor AP location using smartphones. In *Proceedings of ACM MobiCom*, pages 109–120, 2011.
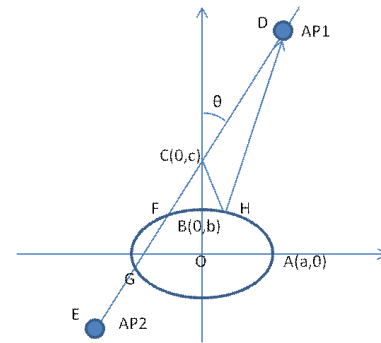
APPENDIX: HUMAN BODY INTERFERENCE ANALYSIS

Fig. 7: Modelling human impact on wireless signals

We take human body as an ellipse, as is shown in Fig. 7. The wireless signals received at AP2 are mainly affected by shadowing. We adopt the logarithmic wireless path model: $P(d) = P_0 - a \log \frac{d}{d_0} + n$. The signal attenuation is much larger within human bodies than in the air. The attenuation is proportional to the relative distance of GF, i.e., $\frac{|CG|}{|CF|}$. With some calculations, we can write the RSSI decrease at AP2 due to shadowing as:

$$\Delta RSSI_2 = \alpha \log \frac{akc + b\sqrt{a^2k^2 - c^2 + b^2}}{akc - b\sqrt{a^2k^2 - c^2 + b^2}} \qquad (2)$$

where $k = ctg(\theta + \pi)$, and α is the coefficient to be empirically evaluated in different scenarios.

On the other hand, the signals received at AP1 is mainly the superposition of the signals on two paths, one line-of-sight path and one reflection path off human body. The affected amount is related to the reflection path length, i.e., $C \rightarrow H \rightarrow D$. We observe that if $|CD| = l$, the distance between C and D is large and $|CB| \approx |BO|$, $|CHD| \approx |DO|$. We can write the RSSI change at AP1 due to reflection as:

$$\Delta RSSI_1 = \beta \log \frac{\sqrt{(l\sin(\theta))^2 + (c + l\cos(\theta))^2}}{l} \qquad (3)$$

where β is the coefficient to be empirically evaluated in different scenarios.