Cycle4Value: A Blockchain-based Reward System to Promote Cycling and Reduce CO_2 Footprint

Alexander K. Seewald[1], Mihai Ghete[4], Thomas Wernbacher[2], Mario Platzer[3], Josefine Schneider[3], Dietmar Hofer[4] and Alexander Pfeiffer[2]

[1]*Seewald Solutions, Lärchenstraße 1, A-4616 Weißkirchen a.d. Traun, Austria*
[2]*Donau Universität Krems, Dr.-Karl-Dorrek-Straße 30, A-3500 Krems, Austria*
[3]*Yverkehrsplanung, Brockmanngasse 55, A-8010 Graz, Austria*
[4]*Bike Citizens Mobile Solutions GmbH, Kinkgasse 7, A-8020 Graz, Austria*

Keywords: Cycling, Mobility of the Future, Blockchain, Reward Model, Cheating Detection, Fraud Detection.

Abstract: In Cycle4Value (C4V), a transparent and low-threshold reward model to promote cycling based on the key technology blockchain is being researched and tested in practice for the first time. The economic, health and ecological benefits are presented in a simple and comprehensible way and, after a plausibility check using a pretrained machine learning model, are converted into a real value, i.e. a cycle token. These units of value are stored in a digital wallet and can be reimbursed in a marketplace set up for testing. The research project goes beyond conventional incentive systems, since 1) the storage of the value units as well as the payment process is decentralised, tamper-proof and transparent, and 2) the real economic and environmental benefit of active cycling is monetized in a fair manner. Initially we describe the background of the project. The main part of this paper concerns ongoing work on the plausibility check which also needs to be able to detect cheating.

1 INTRODUCTION

It is well known that global warming – mainly driven by too high CO_2 concentrations in the atmosphere – is a pressing issue and necessitates significant changes in consumer behaviour over a relative short time period. While legislation may helpful in forcing businesses to change their behaviour, consumer behaviour is much harder to influence directly. One effect that could be easily obtained would be to promote cycling in cities instead of other means of transportation as it is almost completely CO_2 free.

However, despite various measures to promote cycling, the overall proportion of cycling in Austria has improved only slightly in recent years (Tomschy and Steinacher, 2017; Illek and Mayer, 2013). As part of its mobility strategy, the Austrian Federal Government has therefore set itself the target of doubling the proportion of cycling within seven years. To this end, not costly infrastructural measures but motivational or behavioural approaches should be pursued. In this context mobile apps for tracking and incentivisation of one's own data – as a technological manifestation of the *quantified self* – are developing rapidly.

Table 1: Monetary benefits by indicators and action sphere for each bicycling kilometer. All values in € per km.

Action sphere	**Indicators**	**Monetarized savings per bicycle kilometer**	
		per indicator	**per sphere**
Ecology	Reduced CO_2	0.025	0.0377
	Reduced NO_x	0.0105	
	Reduced NH_3	0.0003	
	Red. $NMVOC$	0.00022	
	Reduced SO_2	0.00002	
	Reduced $PM_{2.5}$	0.00006	
	Reduced PM_{10}	0.0009	
Economy	Red. individual mobility costs	0.26	0.26
Health	Health promotion	0.94	1.07
	Improvement of traffic safety	0.13	
Total benefits		**1.37**	

Sweepstakes or even performance-related rewards are paid out depending on the underlying gratification system in the form of cryptocurrencies. Blockchain technology has a significant potential in the handling

Seewald, A., Ghete, M., Wernbacher, T., Platzer, M., Schneider, J., Hofer, D. and Pfeiffer, A.
Cycle4Value: A Blockchain-based Reward System to Promote Cycling and Reduce CO2 Footprint.
DOI: 10.5220/0010318010821089
In *Proceedings of the 13th International Conference on Agents and Artificial Intelligence (ICAART 2021) - Volume 2*, pages 1082-1089
ISBN: 978-989-758-484-8

of user data due to its decentralization, transparency and security (Buhl et al., 2017). Applications on the blockchain are treated as disruptive innovations for a wide range of applications: from transaction processing to land register entries to logistics chains, the middleman is to be cut out in the future, and data is to be stored in a forgery-proof and decentralised manner (Hopf and Picot, 2018). The application of these innovative measures could provide the impetus to promote cycling with all the associated benefits in terms of reduced emissions, positive health effects and reduced infrastructural costs.

The solution envisaged in our project *Cycle4Value* (C4V) will reward cyclists for regular cycling by means of so-called *Cycle Tokens*. As such it is related to other approaches which focus on promoting physical activity through the use of new technologies such as (Noël Racine et al., 2020; Van Hoye et al., 2019; Weatherson et al., 2017). In our case, the technology used in the form of the Ardor Blockchain[1] represents an energy-efficient solution for validating route data and transaction processing. In the sense of a proof of concept, it is to be tested whether and how a safe and transparent process of value generation for regular cycling can be created via utility token, which can translate the macroeconomic effects of cycling into value units.

In the course of a broad field test in Graz, Berlin and Krems, the usability, acceptance and scalability of the developed solution will be analysed. The cost savings at individual and collective level have been evaluated in close cooperation with a stakeholder board and will ultimately be transferred into a market place, which will enable the exchange of cycle tokens into incentives such as discounts or spare parts.

Within this project we have already estimated the cost savings, or benefits, per cycling kilometer for a variety of spheres and indicators – see Table 1. It can be seen that within the ecological sphere, CO_2 has the highest benefits at € 0.025. Health has even larger benefits at € 1.07. These numbers are based on a variety of public data sources and were refined in interviews with domains experts. Since the focus within our research project C4V is on the benefits of cycling, the costs of cycling were neglected. This normally leads to an overestimation of the benefits. However, since all non-monetary indicators that have an additional benefit are not considered, it can still be assumed that the monetary end values are rather conservative. All calculations are based on the simplifying assumption that each bicycle ride represents a saved car ride. While this simplification does not do justice to the complexity of reality, it still represents a common scientific approach to calculating costs or benefits.

The remainder of this paper focusses on the ongoing work within this project concerning plausibility and cheating detection to ensure that cycle tokens are only obtained via live bike rides.

[1] https://www.jelurida.com/ardor

2 RELATED RESEARCH

As we found out from extensive literature research, the task of identifying fraud or finding cheaters in the bike sector (i.e. those tracks that do not originate from live bike rides) is completely new. Only the well-examined area of Transportation Mode Detection (TMD) is related, as cheaters sometimes use other means of transportation than bicycles.[2] In later experiments we will show that pre-trained systems for TMD show competitive results compared to specifically trained systems and can reconstruct plausibility according to a legacy system about equally well, but without having seen the GPS tracks to be classified before.

2.1 Motion Detection with 2D/3D GPS Tracking

This group of systems detects means of transport using 2D (without altitude) or 3D (with altitude) GPS position data. Since up to now we only had 3D GPS data available (see Table 2), these are the only systems that could be tested directly.

(Dabiri et al., 2019) describes a state-of-the-art deep learning system for motion detection based on 2D GPS data. Both classic and DL-trained features are used, but the difference is relatively small (see also (Etemad, 2018)). The GeoLife 1.3 dataset was used for training. Only transportation modes for walking, cycling, bus, car/taxi and subway/railway/expressway were trained. The GPS tracks were divided into segments when breaks were taken or a maximum length was reached. An overall accuracy of 76.8% and F_1 (i.e. the balanced F-measure) of 0.764 was achieved. A thesis by the same author gives more precise details as well as the complete training code including the data used for the training, which enabled us to reconstruct the pre-trained system exactly. However, the

[2] In Section 3, some attacks are mentioned where this is not the case, and which could not be detected by a TMD system, notably 1, 4 and 6.

[3] All plausibility values were recomputed with latest legacy model.

[4] See (Zheng et al., 2010; Zheng et al., 2009).

Table 2: Datasets used within this paper. All contain GPS-3D data (longitude, latitude, altitude, timestamp).

Name	#GPS tracks	#GPS points	Class	Created by	Publicly available
BC-GPS-10k	10,000	10,861,318	Plausibility	Bike Citizens Mobile Solutions	No
BC-GPS-4M	4,082,450	4,439,163,525	Plausibility	Bike Citizens Mobile Solutions	No
BC-GPS-4M-RCP[3]	4,082,450	4,439,163,525	Plausibility	Bike Citizens Mobile Solutions	No
GeoLife 1.3[4]	17,621	24,876,978	Transport Mode (on subset of data)	Microsoft	Yes

written code was not scalable on several levels – runtime, main memory and hard disk space requirements – and required extensive modifications to process our BC-GPS-4M dataset.

(Etemad, 2018) describes a PhD thesis that uses classical learning algorithms (Random Forest, Support Vector Machines, Decision Trees, Multi-Layer Perceptron, AdaBoost, ...) and handcrafted features as well as other preprocessing steps. In his section 4.4. it is shown – among other results – that the obtained performance is significantly better than the one described in (Dabiri et al., 2019), but only at a significance level of p=0.0796 (92.04%) which is less stringent than the standard 95% level normally used. This shows that in this field Deep Learning algorithms are not yet superior and that classical learning algorithms can still obtain equivalent results.

(Nawaz et al., 2020) describes a Long-Short-Term-Memory (LSTM) Deep Learning model for motion detection. It was again trained on the GeoLife dataset as in (Dabiri et al., 2019; Etemad, 2018), but unlike there, the class subway/railway/train was not considered here. The results (Accuracy=83.81% and F_1=0.8397) are therefore not directly comparable and are in fact slightly better than the other papers – possibly because of the smaller number of classes. Again, classical features were calculated (i.e. speed, acceleration, and other derivatives of position) and the GPS tracks were divided into individual segments. To take advantage of the good performance of Deep Learning networks on image data, the tracks were then scaled and mapped to a 2D image and provided as additional input. For the high effort involved, however, the results were not better than much simpler methods. For example, the LSTM model already achieves an accuracy of 79.15% without these 2D image features and the results in (Dabiri et al., 2019; Etemad, 2018) are not much worse.

In general, it was surprising that - contrary to the use of deep learning networks in speech, image and video processing - no competitive Deep Learning system yet exists that can process GPS data directly without preprocessing. Also, (Dabiri et al., 2019) uses mostly classical preprocessing, and the most important features identified are relative distance, speed and acceleration, which are also often used in classical systems. As expected, learning algorithms are universally used.

2.2 Motion Detection with Local Motion Sensors

(Soares et al., 2019) compares numerous local Android mobile phone sensors (including some derived from multiple sensors such as Orientation), classical and deep learning learning algorithms on the relatively small TM dataset. As pre-processing, common time-based features (max, min, entropy, average, variance, median, standard deviation, quartile, ...) and FFT features (spectrum, highest amplitude, spectral density and entropy, ...) were calculated on the raw data of all sensors. The used locomotion classes were standstill, car, walking, railway and bus.[5] The propagated deep learning model TMD-LSTM is not the best model (Acc=90%, F_1=0.90), but with 314 KB it is relatively small. However, the Decision Tree model is half as large and similarly good (Acc=91%, F_1=0.87) and certainly has comparable computational complexity.

(Vu et al., 2016) describes one of the few models that can process sensor data directly without preprocessing. However, it is limited to accelerometer sensors. Here, the somewhat older HTC dataset is used, from which the classes standstill, walking, running, bicycle, and other means of transport (a combination of motorbike, car, bus, underground, train) are trained. The sensor data is divided into windows of 12s length, but the system can also predict a class every 165ms after the training with a few tricks.[6] A

[5] Sadly, cycling is missing.

[6] A runtime of approx. 28ms per prediction is given, albeit without information about the computing platform.

new type of recurrent network, CGRNN, is presented and trained. The result is an overall window-level accuracy of 93.10% (90.93% for bicycle only), which is a very good result for a single sensor without any explict preprocessing.

There is still a lot of research in the related field of activity detection using similar sensors, preprocessing steps and models, but cycling is not normally included as a separate activity. For this reason, this kind of work is not considered here.

Finally, it can be seen that of the local sensors, the accelerometer seems to be the most important. Since the gyroscope is already universally used and can be read out easily with little energy consumption, we will also consider it. With the derived sensors it should be noted that although they save some computational effort (e.g. for Orientation, which calculates the direct orientation of the smartphone to Earth – thanks to Gimbal Lock and other special cases a relatively complex algorithm), they do not necessarily exist universally in the same way on all platforms. So a review of the existing APIs (interfaces) would certainly be a first useful step.

2.3 Combined Models for Motion Detection

(Nitsche et al., 2014) describes a motion detection system using GPS and accelerometers as a complementary source of information for travel surveys. The classes used were walking, bicycle, motorbike, car, bus, tram, metro, rapid transit and standstill. A separate test set of 2,089 trips (of which about 50% were walking and 10% were cycling) was recorded and used for training and evaluation. As learning algorithm an ensemble learning algorithm and a DHMM were used to reconstruct the class and additional transition probabilities. The pre-processing of the sensor data is similar to (Soares et al., 2019) (i.e. time- and FFT-based features in a fixed window). With leave-one-out cross-validation, precision for bicycle was 0.88, recall 0.95 (F_1=0.91) and accuracy 95.95%, which is slightly better than in (Vu et al., 2016). The overall accuracy was 80% across all classes, but for more and more similar classes than in other systems.

(Feng and Timmermans, 2013) compare self-trained systems with GPS, accelerometers and combined data. The combined data gave the best results. A number of time-based features were used as pre-processing, including derivations from altitude according to GPS (i.e. 3D GPS), which is welcome as so many papers ignore GPS altitude. A Bayesian Belief Network was used as the learning algorithm. The classes used were *Activity, Walking, Running, Bicycle, Bus, Motorcycle, Car, Train, Underground, Tram* and *Express Train*. A dataset with about 81,000 entries (recorded on identical proprietary devices) was used. The combined model had detection rates between 83% (activity, train) and 100% (walking, bicycle, motorbike). For bicycle detection, however, both the GPS-only and the combined model were already perfect at 100% and the accelerometer model was significantly worse at 88%. However, since unrealistic data was manually removed and the class *Activity* is somewhat unclearly defined, the values mentioned may be a bit too high. With the original data, the bicycle detection with 95% accuracy in the combined model and 94% in the accelerometer-only model shows much less difference, which seems more plausible to us.

2.4 Alternative Models

(Charvátová et al., 2017) describes a system for analysing the pulse rate when cycling using GPS. A specific cycle route was followed 41 times and the recorded 3D GPS data was related to the pulse rate. Different classical ML models were trained to differentiate between up and down cycling classes. This could be seen both from the altitude gradients of the GPS data and from the pulse rate (with a time delay) with accuracies between 93% and 97%. Although the objective was different and had a medical background, the clearly recognisable patterns of pulse rate in cycling would make cheating much more difficult. In addition, information about the health status of the user would be directly obtainable. However, the use of additional sensors is not feasible within our project.[7]

2.5 Blockchain

(Thomas et al., 2019) investigate various blockchain systems with regard to their applicability for storing personal metadata for grades and certificates. The requirements for the blockchain system are similar to those within our project C4V. A public blockchain system is used, where utility-tokens are generated and sent under certain conditions. These preconditions are for example that tokens cannot be forwarded without the operator's consent, the separation of private

[7]It would be possible to use the front camera of the mobile phone – if the holder is positioned appropriately – to determine the pulse rate from the video data (pulse-by-face) and thus measure it for such a model. However, this idea is out-of-scope for this project regarding the complexity of such an implementation, and it would also increase power consumption of the smartphone app quite dramatically.

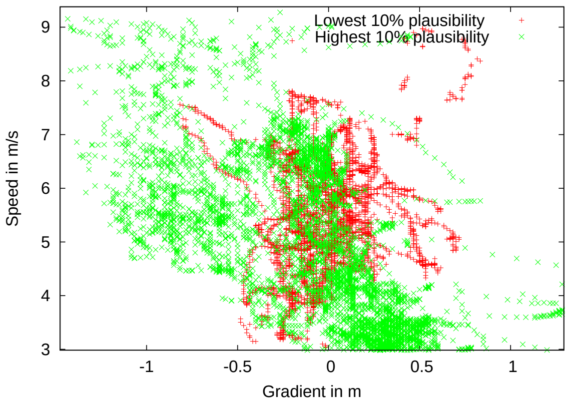

Figure 1: Scatterplot of gradient vs. speed, colored by 10% lowest and highest plausibility.

and public information, the assumption of transaction costs by the operator, the possibility to operate an own node on the own laptop or smartphone and even the possibility that private data may be automatically deleted from the blockchain after a fixed period of time with only the evidence of the transaction itself remaining.

3 POSSIBLE ATTACKS

The semi-monetary compensation in form of tokens which are potentially convertable[8] to cash makes cheating the system (respectively fraud) much more likely. While catching cheaters is essential to create and maintain trust in the reward system, and therefore false negatives (Type II errors, i.e. not detected cheaters) should be small, the number of false positives (Type I errors, i.e. people wrongly classified as cheaters) should also be small to prevent disillusionment and reduced trust in the reward system. To some extent this can be optimized by using learning algorithmus that output confidence values and choosing appropriate thresholds.

The number of tokens generated corresponds to the square root of cycling kilometers per track, and is cut-off at two levels: at most 4 tokens per track, and at most 8 tokens per day, yielding some buffer against attacks by limiting payoff for a single user. However as this is probably not sufficient, we have collected six types of possible attacks on the system. This list may be of course be incomplete and is only intended as a first overview.

1. Uploading tracks from sports events (e.g. cycle racing)
2. Uploading tracks from delivery services
3. Uploading tracks made with other vehicles (may include e-Bikes)
4. Using more than one phone on a bicycle
5. Uploading automatically generated fake tracks
6. Uploading modified real-life tracks

Attacks 1,2 and 3 are not a large problem. In 1, the restriction of the maximum number of tokens per track and day – mentioned above – severely restrict the obtainable gain per user. 2. should not yield any payoff as these bike rides would happen anyway even without incentives. However as long as only a single person profits, it is again not a large problem. We could address it by either filtering out common bike routes for delivery services, or severely reducing their payoff. The same approach could be applied to attack 1. should too many similar tracks be uploaded. Attack 3 should be easily detectable by the fraud detection system although the differentiation of e-bikes is still an open problem.[9]

The more interesting cases are attacks 4, 5 and 6. 4. should be detectable – especially once we have gyroscope and accelerometer sensors – by aligning multiple tracks and determining their similarity. The local sensor movements on these tracks will be extremely similar, which however can only be detected on the level of multiple tracks and not from a single track. We will address this in future work.

For attack 5., it is important that not all details of the system are known, otherwise engineering a reverse system is made correspondingly easier. However, since we focus on machine learning techniques throughout, it is likely that creating realistic fake tracks will be very hard, and people will mainly resort to replaying modified existing tracks (6.), which should again be detectable by their similarity. Also, only 4-6 (and 3 when combined with 4) ensure potentially unlimited payoff, which is surely the most tempting fraud scenario. It should be noted that only attacks 3 and 5 can be addressed at single track level with models described here. The other four attacks must be addressed at the level of multiple tracks and will need completely different learning systems.

As this list may be incomplete, we plan to start a competition for cheating with appropriate incentives before deploying the system widely. This should give us sufficient data on other scenarios not considered here.

[8]E.g. by reselling obtained discounts and spare parts.

[9]We speculate that gyroscope and acceleration data could be used to determine approximate bicycle mass on sharp turns, however it would still not allow us to differentiate e-Bikes from Styrian "Waffenräder" (an old type of bike which weights about 20 kilograms, similar to a modern e-Bike).

Table 3: Experimental results concerning replication of the legacy model. Precision and recall of minority class *no* (plaus.<50).

Dataset	Learning Algorithm	Evaluiation type	Acc.	Prec.	Rec.	F_1	AUC
BC-GPS-10k	JRip/RIPPER	10-fold CV	79.67%	0.772	0.776	0.774	0.814
BC-GPS-4M	JRip/RIPPER	test set (trained on BC-GPS-10k)	83.04%	0.432	0.779	0.556	0.824
BC-GPS-4M	TMD	test set (trained on GeoLife1.3)	81.20%	0.950	0.843	0.893	n/a
BC-GPS-4M-RCP	TMD	test set (trained on GeoLife 1.3)	72.02%	0.764	0.877	0.817	n/a

4 EXPERIMENTS

4.1 Legacy Model

At the start of our project, BikeCitizens had already deployed a plausibility checking system since 2013, primarily to prevent the upload of obvious non-cycling tracks. It outputs a plausibility of 0 to 100 (in integer values) with 100 being the highest possible confidence in the given data being a real cycling track, and 0 the lowest. The algorithm was mainly tracking speed values and giving positive points for plausible and negative points for implausible speeds. These points were summed up separately, thresholded to manually determined maximum values, weighted by track length and normalized to obtain the plausibility value. Since 2013, the algorithm was changed several times in minor ways,

While it is a legacy system and does not correspond exactly to our stated purpose to build a cheating and fraud detection system, it still remains a resonable starting point for a first analysis.

Initially – for the rule learning model – we worked with BC-GPS-10K (see Table 2), but later we received a much larger dataset, BC-GPS-4M, which was used for subsequent experiments.

4.2 Rule Learning Model

One somewhat trivial observation is that when cycling, going downhill is usually much faster than going uphill. If we therefore compute speed and gradient and plot them against each other, coloring by the lowest and highest 10% plausibility values – see Fig. 1 – we can distinguish this pattern quite clearly. The fake tracks mostly show a pattern without any dependence on gradient except for very few outliers, and in some cases the gradient is even opposite to what we would expect (i.e. higher speed for higher gradient).

Emboldened by these results, we aimed to replicate the ad-hoc model using the rule learning algorithm *JRip*, a java-based reimplementation of RIPPER by (Cohen, 1995). We first discretized the plausibility value into two approximately equally sized bins: from 0-49 (no) and the value 100 (yes) on its own. As input we discretized the gradient (computed between each sample, i.e. once a second[10]) into four[11] bins: $(-\infty,-1)$, $[-1,0)$, $[0,1)$ and $[1,+\infty)$ For each bin, we then computed mean ($\bar{x}$), standard deviation (σ) and relative standard deviations ($\frac{\sigma}{\bar{x}}$) of 3D speed according to consecutive GPS coordinates (including altitude). Additionally, we added normalized mean values over all bins ($\forall_i\{\frac{\bar{x}_i}{\sum_{j=1}^{4}\bar{x}_j}\}$).

By ten-fold crossvalidation on all selected tracks from BC-GPS-10k we obtained a precision of 0.772, recall of 0.776, F_1 of 0.774 and area-under-ROC 0.814 for class *no* (i.e. low plausibility). All results can also be found in Table 3. An analysis of the model from the whole dataset shows that two out of the four rules obtains use a lower bound for downhill cycling (corresponding to *mean0/1*) and an upper bound for uphill cycling (corresponding to *mean2/3*). This clearly indicates that the model found and utilized the somewhat trivial pattern mentioned earlier.

After obtaining the larger BC-GPS-4M dataset, we applied the previously trained model to it and obtained somewhat worse results: a precision of 0.432, recall of 0.779, F_1 of 0.556 and area-under-ROC of 0.824. This may have been caused by the inconsistently computed plausibility measure in this larger dataset, as noted later.

4.3 TMD Model

As already mentioned, only (Dabiri et al., 2019; Etemad, 2018; Nawaz et al., 2020) could be evaluated because we only received 3D-GPS data for evaluation. Since (Etemad, 2018) is more of a classical

[10] We also tested 5s and 15s – results were slightly worse.
[11] We also tried two and three bins centered on zero. Here, the results were very similar.

model and (Nawaz et al., 2020) presents a very complex model that would be very time-consuming to implement from scratch, we chose (Dabiri et al., 2019) where the underlying code is also available in a usable form in the corresponding master thesis. This however does not include the pre-trained Transport Mode Detection (TMD) model itself, which we had to train ourselves on the GeoLife 1.3 dataset.

We were interested to find how well the pre-trained model would be able to recontruct plausibility without any retraining at all. We therefore computed the proportion of segments from each track in BC-GPS-4M with transport mode classification *cycling*, weighted the results by model confidence and normalized it to $[0,100]$, thus creating an alternative plausibility measure. On the original BC-GPS-4M data set, the results of this pre-trained model with accuracy 81.20% and F_1=0.893 are even slightly better than the classical motion detection – and this without any actual training.

Since there was some concern that the plausibility measure may not have been consistent over the whole dataset as it was stored at the time of recording and the code was sometimes changed, we recomputed the plausibility measure using the most recent code and obtained BC-GPS-4M-RCP, which was used to all remaining experiments. When recomputing performance of the TMD model with the recomputed plausibility values, accuracy is reduced to 72.02% and F_1 to 0.817. However as we will see this may have been caused by a shift in the plausibility values that may be easily addressable by recalibrating the TMD model.

To better compare both measures, we split the original plausibility values into five (almost) equally large intervals: $[0,20),[20,40),[40,60),[60,80),[80,100]$ and computed arithmetic mean and standard deviation of the new plausibility measure over each interval. Fig 2 shows the results and also the same values averaged for each possible legacy model plausibility as *Raw Mean*.[12] The latter roughly corresponds to a ROC curve. As you can see, the mean TMD-derived plausibility measure increases strictly monotonically from the lowest to the highest interval. Although standard deviation is initially high, it shrinks for higher intervals. A recalibration of the TMD model or an adaptation of the threshold for plausibility could improve the match considerably. The Pearson's correlation coefficient between legacy and TMD measure is 0.31, indicating a weak to moderate correlation. Note again that the TMD model has **not** been trained on **any** part of the BC-GPS datasets,

[12]This corresponds to using 101 bins, one for each plausibility values in the interval [0,100].

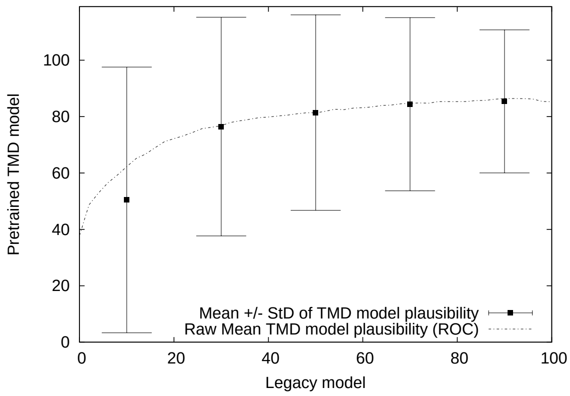

Figure 2: Comparison of the ad-hoc plausibility (legacy) model with transport-mode-detection (TMD) model.

so these results are quite surprising and indicate that TMD models are a good starting point for a trainable plausibility – and possibly also cheating – detection system. Note also that the legacy model was not explicitly designed to detect cheating and fraud, but rather to increase the quality of uploaded data by ensuring that only bicycle tracks were accepted.

5 CONCLUSION

We have shown that the results of a legacy model for plausibility detection of cycling tracks can to a moderate extent be reconstructed by a classically trained rule learning model with handcrafted features, as well as to a somewhat lesser extent by a pretrained transport mode detection (TMD) model, even though the latter was trained on a completely different dataset. Improvements to the TMD model may be made by adding GPS altitude to its input data and model structure; retraining at least part of the existing plausibility data; and calibrating the model to give more similarly distributed plausibility outputs.

However, as our primary goal is not to replicate the existing legacy model but rather to build a more general model both for plausibility and cheating/fraud detection, more data is needed. To address this and other open issues, we will 1) Start a competition for cheating with appropriate incentives for a small test groups; 2) Research fast algorithms for track alignment, thus taking care of *Attack 4: Multiple phones on one bike*; 3) After deployment, regularly analyze outliers with high and/or similar payoff and update the system accordingly, keeping in mind the trade-off between Type I and II errors.

We will also consider the use of synthetically generated data for various attack types to expand our datasets. As the most tempting attacks call for a large number of users to be automatically created, another option may be to simply make user registration non-scriptable e.g. by using captchas or requiring a short cycling sequence with pulse rate (for example, by using the front smartphone camera and pulse-by-face).

ACKNOWLEDGEMENTS

This project is currently funded by the Austrian Research Promotion Agency (FFG) and by the Austrian Federal Ministry for Climate action, Environment, Energy, Mobility, Innovation and Technology (BMK) as project Cycle4Value (873384).

REFERENCES

Buhl, H. U., Schweizer, A., and Urbach, N. (2017). Blockchain-Technologie als Schlüssel für die Zukunft. *Zeitschrift für das gesamte Kreditwesen*, pages 596–599.

Charvátová, H., Procházka, A., Vaseghi, S., Vyšata, O., and Vališ, M. (2017). GPS-based analysis of physical activities using positioning and heart rate cycling data. *Signal, Image and Video Processing*, 11(2):251–258.

Cohen, W. (1995). Fast effective rule induction. In *Proceedings of the Twelfth International Conference on Machine Learning, San Francisco, CA, 1995*, pages 115–123. Morgan Kaufmann.

Dabiri, S., Lu, C.-T., Heaslip, K., and Reddy, C. K. (2019). Semi-supervised deep learning approach for transportation mode identification using GPS trajectory data. *IEEE Transactions on Knowledge and Data Engineering*, 32(5):1010–1023.

Etemad, M. (2018). Transportation Modes Classification Using Feature Engineering. *arXiv preprint arXiv:1807.10876*.

Feng, T. and Timmermans, H. J. (2013). Transportation mode recognition using GPS and accelerometer data. *Transportation Research Part C: Emerging Technologies*, 37:118–130.

Hopf, S. and Picot, A. (2018). Revolutioniert Blockchain-Technologie das Management von Eigentumsrechten und Transaktionskosten? In *Interdisziplinäre Perspektiven zur Zukunft der Wertschöpfung*, pages 109–119. Springer.

Illek, G. and Mayer, I. (2013). Radverkehr in Zahlen–Daten, Fakten und Stimmungen. *BMVIT, Wien*.

Nawaz, A., Zhiqiu, H., Senzhang, W., Hussain, Y., Khan, I., and Khan, Z. (2020). Convolutional LSTM based transportation mode learning from raw GPS trajectories. *IET Intelligent Transport Systems*, 14(6):570–577.

Nitsche, P., Widhalm, P., Breuss, S., Brändle, N., and Maurer, P. (2014). Supporting large-scale travel surveys with smartphones–A practical approach. *Transportation Research Part C: Emerging Technologies*, 43:212–221.

Noël Racine, A., Garbarino, J.-M., Corrion, K., D'Arripe-Longueville, F., Massiera, B., and Vuillemin, A. (2020). Perceptions of barriers and levers of health-enhancing physical activity policies in mid-size french municipalities. *Health research policy and systems*, 18:1–10.

Soares, E. F. d. S., Salehinejad, H., Campos, C. A. V., and Valaee, S. (2019). Recurrent Neural Networks for Online Travel Mode Detection. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE.

Thomas, A., Koenig, N., Higgins, T., Black, M., Pfeiffer, A., Donelan, L., Lenzen, B., Muniz, N., Patel, K., Taylan, A., and Wernbacher, T. (2019). *From Learning to Assessment, How to Utilize Blockchain Technologies in Gaming Environments to Secure Learning Outcomes and Test Results*, page 172. Malta College of Arts, Science and Technology.

Tomschy, R. and Steinacher, I. (2017). Österreich unterwegs… mit dem Fahrrad. *BUNDESMINISTERIUM FÜR VERKEHR, IUT, WIEN (ed.)*.

Van Hoye, A., Vandoorne, C., Absil, G., Lecomte, F., Fallon, C., Lombrail, P., and Vuillemin, A. (2019). Health enhancing physical activity in all policies? comparison of national public actors between france and belgium. *Health Policy*, 123(3):327–332.

Vu, T. H., Dung, L., and Wang, J.-C. (2016). Transportation mode detection on mobile devices using recurrent nets. In *Proceedings of the 24th ACM international conference on Multimedia*, pages 392–396.

Weatherson, K. A., McKay, R., Gainforth, H. L., and Jung, M. E. (2017). Barriers and facilitators to the implementation of a school-based physical activity policy in canada: application of the theoretical domains framework. *BMC public health*, 17(1):835.

Zheng, Y., Xie, X., Ma, W.-Y., et al. (2010). GeoLife: A collaborative social networking service among user, location and trajectory. *IEEE Data Eng. Bull.*, 33(2):32–39.

Zheng, Y., Zhang, L., Xie, X., and Ma, W.-Y. (2009). Mining interesting locations and travel sequences from GPS trajectories. In *Proceedings of the 18th international conference on World wide web*, pages 791–800.