Social network analysis of electric vehicles adoption: a data-based approach

V. Breschi\textsuperscript{a}, M. Tanelli\textsuperscript{a,b}, C. Ravazzi\textsuperscript{b}, S. Strada\textsuperscript{a} and F. Dabbene\textsuperscript{b}

Abstract—Mobility is undergoing dramatic transformations. Especially in the context of urban areas, several significant changes are underway, driven by both new mobility needs and environmental concerns. The most mature one, which still is struggling to affirm itself is the process of the adoption of Electric Vehicles (EVs), thus switching from fuel-based to battery-powered propulsion technologies. Many social and economic barriers have proved to play a crucial role in this process, ranging from level of education, environmental awareness, age and census. This work aims at contributing to the study of this adoption process through a data-based lens, using real mobility patterns to setup a social-network analysis to model the spread of consensus among neighboring people that can enable the switch to EVs. In particular, we build the network topology using proximity measures that emerge from the analysis of real trips, and the initial disposition of the single agents towards the EV technology is inferred from their real mobility patterns. Based on this network, a cascade adoption model is simulated to investigate the dynamics of the adoption process, and an incentive scheme is designed to show how different policies can contribute to the opinion diffusion over time on the network.

I. INTRODUCTION

The way in which people are moving, especially in urban areas, is experiencing several major paradigm shifts. These are necessary to answer the needs of a growing number of people living that are aggregating to live in the so-called mega-cities (more that 60% of the world population is expected to leave in a mega-city by 2030), while being sustainable and energy efficient, [1].

Two of the main directions of change that mobility is facing are the following: (i) from internal-combustion-engine (ICE) to battery-operated electric vehicles (BEV); oil-based propulsion needs to be abandoned and substituted by fully electric or at most hybrid propulsion for long-range vehicles. In fact, transportation in its current form is responsible for the emission of over a quarter of all greenhouse gases (GHG) in Europe; in particular, passenger cars alone account for 41% of such emissions (11% of the total), [2]. A shift towards more eco-friendly means of mobility such as EVs could substantially contribute to lowering the amount of GHG emissions caused by road transport. This holds especially true when we consider that EVs are able to convert approximately 60% of the electrical energy from the grid to power at the wheels, while conventional ICE vehicles only convert approximately 20% of the fuel energy to power at the wheels, [3], [4]; (ii) a transition from ownership to usership: vehicle-sharing systems seem one of the most promising way for sustainable mobility in urban areas. With more than half of the worlds population that will soon live in metropolitan areas, the current private-vehicle-based mobility would lead to unmanageable traffic density, energy consumption, pollution, and congestion. Sustainable and shared mobility solutions in the context of highly-densely populated areas can also help to fulfill the mobility needs of medium to low-income people, as they lower the ownership costs of an EV.

From the above discussion, it is clear that mobility as a whole will be more and more developed intertwining technological, social and economic needs. As such, the new, sustainable and shared mobility models cannot but be designed with human-centered principles, so as to embed in their functioning interactions with the users, and to manage their adaptation and evolution not only on the basis of technological needs, but also of social ones, [5]. In the two examples above, a fundamental mindset shift must take place in the final user to allow a significant penetration of the two new mobility models, i.e., electric and shared vehicles. In fact, as far as EVs are concerned, despite the sales growth of the recent years (registering a year-on-year market growth of 63% in 2018, [6]) the penetration of electric cars is still limited to less than 1% of the global car fleet, [6]. The greatest barriers to widespread EV adoption are the purchase price, the driving range and the charging availability, [7]. Even though new EVs have considerable driving ranges, still the so-called range anxiety, i.e., the fear that users have to remain without battery charge while still on the road, is still a strong inhibiting factor in the EV diffusion. Also in the shift from ownership to usership strong social changes must be in place. Up to now owning a (possibly high-end) car is a strong social marker associated to high-status of individuals. Removing such social link to ownership is one of the main keys to allow the spread of the new model. A sensible framework to mix technological and social variables for analyzing the aforementioned adoption processes is that of opinion dynamics, which allows modelling the formation of beliefs, the aggregation of opinions and their dynamic evolution in a hyper-connected society, all within a quantitative and rigorous framework, see e.g., [8], [9], [10], [11] and references therein.

In this work, we concentrate on the analysis of the EV-adoption process, and we use as a basis for modelling the mobility habits a large data-set of vehicle-based measurements, obtained via telematic e-Boxes, that followed thousands of

---

\textsuperscript{a}) Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy.
\textsuperscript{b}) Istituto di Elettronica e Ingegneria dell’Informazione e delle Telecomunicazioni - IEIIT CNR Corso Duca degli Abruzzi 24, 10129 Torino.
users in a large Italian province for one year in their daily trips. These data are used to extract real mobility patterns that quantitatively describe the feasibility, for each single vehicle, to be seamlessly replaced by an EV. Based on this data-based scenario, we build a social network description of the considered users based on a proximity measure, and then analyze the EV-adoption process under different perspectives. Further, we show how incentives policies can sustain and speed this adoption process. To the best of our knowledge, this is one of the very first contribution that uses the opinion dynamic framework in a data-driven way, which we believe is essential to consider the impact of social variables while starting from a realistic picture of the current needs of todays mobility. We believe that only a meaningful mix of the two can yield sensible answers that allow taking into account users characteristics in the design of new mobility services.

The structure of the paper is as follows. Section II introduces the graph-based model based on which the network topology is defined. Then, Section III illustrates the data-based analysis of the EV-adoption process, and how from that analysis we extracted the needed information for setting up the social network within an opinion dynamics framework. Based on it, Section IV analyzes the diffusion of EV adoption under different perspectives, and Section V discusses the effects of possible incentive policies to support a faster adoption spread.

II. CASCADING NETWORK MODEL

For modelling the EV-adoption process of interest, we need to build the influence network that describes the interaction between users based on proximity. To do that, we will now describe the cascading network model that will be the basis for the subsequent analysis. To start with, let us consider a simple and undirected graph \( G = (\mathcal{V}, \mathcal{E}) \) where \( \mathcal{V} \) is the set of nodes representing the agents and \( \mathcal{E} \subseteq \mathcal{V} \times \mathcal{V} \) is the set of edges. The assertion \((u, v) \in \mathcal{E}\) implies that agent \(v\) is influenced by agent \(u\) and, since the graph is undirected, also \((v, u) \in \mathcal{E}\). We denote the set of in-neighbors of a node \( v \in \mathcal{V} \) by \( \mathcal{N}_v = \{ w \in \mathcal{V} : (w, v) \in \mathcal{E} \} \). A graph \( G \) is called strongly connected if there is a path from each vertex in the graph to every other vertex.

We consider a deterministic cascade model on network \( G \). Each node is endowed with a threshold value \( \alpha_v \in [0, 1] \). At time \( t = 0 \) a subset of nodes \( S_0 \) is chosen to be the seed set, representing the initial EV-adopters. Letting \( S_t \) be the set of new EV-adopters at time \( t \) and the set \( S'_t := \bigcup_{\tau=0}^t S_\tau \) represent the total amount EV-adopters in the network until time \( t \), for each discrete time \( t \geq 1 \) one has

\[
S_t = \left\{ v \in \mathcal{V} \setminus \bigcup_{\tau=0}^{t-1} S_\tau : \frac{|S'_t \cap \mathcal{N}_v|}{|\mathcal{N}_v|} \geq \alpha_v \right\}.
\]

The dynamics in (1) describes a model of EV diffusion where the adoption of the agents depends on the relative popularity of EVs among neighbors.

It should be noticed that \( S'_t \) is monotonically increasing in \( t \in \mathbb{N} \) and the dynamics converge to a final adopters set. In [8] the final adopter set is characterized in terms of the network, seed set and threshold values making use of the concept of cohesive sets, the definition of which is as follows.

**Definition 1 (Cohesive set):** A set \( \Omega \subset \mathcal{V} \) is said cohesive if for all \( \omega \in \Omega \)

\[
\frac{|\Omega \cap \mathcal{N}_\omega|}{|\mathcal{N}_\omega|} > 1 - \alpha_\omega.
\]

This definition states that a set \( \Omega \) is cohesive if for each element \( \omega \) in the set the ratio of neighbors which do not belong \( \Omega \) is strictly smaller than the value of the threshold \( \alpha_\omega \).

**Theorem 1 (Lemma 2 in [8]):** Given a network with seed set \( S_0 \subset \mathcal{V} \), let \( \Omega \subset \mathcal{V} \setminus S_0 \) be the cohesive set with maximal cardinality. Then the dynamics converges in finite time, and the set of final adopters is given by \( S^* = \mathcal{V} \setminus \Omega \).

Computing the maximal cohesive set contained in the complement of the seed set is computationally expensive and an algorithm, based on linear programming methods, is proposed in [12].

It is worth remarking that our goal is to study the spread of adoption in a finite time horizon. More precisely, in next section the cascade EV-adoption model is simulated for \( t \in [0, T] \), using a network topology and features of the model that are directly inferred by real mobility patterns, as will be better explained in the next section.

III. DATA-BASED DESCRIPTION OF THE EV ADOPTION PROCESS

To build the network described before, we use a set comprising data of 1,000 vehicles registered in the Italian province of Parma. The available dataset includes anonymized GPS traces collected over one year (from 1st September 2017 to 31st August 2018) from conventional ICE vehicles equipped with telematic e-Boxes. Each record contains information on the instantaneous GPS latitude and longitude related to an event (such as vehicle ignition or shutdown), along with their associated time stamp. By properly aggregating the available information, trips and stops can be extracted from these records, along with their duration and the traveled distance, thus allowing us to infer the mobility pattern of each vehicle. The reader is referred to [13] for further details on the dataset.

When studying the EV-adoption process, it seems reasonable to assume that one’s opinion can be influenced by the point of views of its closest neighbors. This models that fact that actually seeing neighbors buying new EVs can help fostering individuals’ inclination to do a similar choice. In this light, EV adoption is investigated based on the proximity between the homes of the agents, represented by those stops that take place mostly overnight and that are sufficiently long. For each agent, the GPS coordinates of its home are retrieved from the corresponding records as the centroids of the clusters of proximal stops that contain at least 50 % of the overnight stops. In turn, the latter are empirically detected as

1The 1,000 vehicles used in this study are randomly selected among the 21,251 available ones, so to maintain the statistics on EV suitability presented in [13].
long stops (duration above 7 hours) that take place (at least partially) during night hours.

More formally, \( h_v = (\text{long}_v, \text{lat}_v) \) be the GPS coordinate associated to the home of the \( v \)-th agent, with \( v \in \mathcal{V} \), and denote as \( d(h_v, h_w) \) [km] the geodesic distance between the homes of the \( v \)-th and \( w \)-th agents. For the proximity-based study to be well grounded, the initial dataset is prepossessed to exclude all vehicles \( v \in \mathcal{V} \) such that

\[
d(h_v, h_w) > 60 \text{ [km]}, \quad \forall \, w \in \mathcal{V}, \, w \neq v,
\]

namely we neglect the agents whose home is not sufficiently close to the one of all the others. This leads to a minor reduction in the number of agents considered in our study, with the preprocessed dataset comprising the records of 994 vehicles.

A. Network extraction

Let us assume that the homes of neighboring agents are proximal if they are separated by less than \( d_{\max} = 5 \) [km]. We stress here that, in extracting the data-driven network to investigate EV adoption, we aim at obtaining a connected graph.

By following the rationale dictated by home proximity, we construct a preliminary network \( G = (\mathcal{V}, \mathcal{E}) \), such that

\[
(v, w) \in \mathcal{E} \iff d(h_v, h_w) \leq 5 \text{ [km]}.
\]

However, the graph \( G \) does not result to be connected, thus entailing the presence of groups of agents with intra-cluster distance verifying the condition in (3), but which are actually detached from the other nodes. Since we want to obtain a connected network, we prune \( G \) by progressively removing the agents that fulfill the following conditions.

- **Remote agents**, whose home is separated from all the others by more than \( d_{\max} \) km, namely \( v \in \mathcal{V} \) such that
  \[
  d(h_v, h_w) > 5 \text{ [km]}, \quad \forall \, w \in \mathcal{V}, \, w \neq v.
  \]

B. Data-driven agents classification

To select the seeds \( S_0 \), representing those agents who are assumed to actually having switched to an EV already, we rely on information on the electrification potential of each agent to EV adoption based on the inferred individual motion patterns, by exploiting the rationale in [13]. Accordingly, the suitability to EV adoption is defined based on the characteristics of the daily trips performed by each agent, whenever the corresponding vehicle is used. In particular, the agents are classified according to the distance covered each day, under reasonable assumptions on both the driver’s behavior and the technological limitations of EVs, e.g., the EV range (without recharge) is fixed at 200 km, see again [13].

By restricting ourselves to two classes only (adopters and non-adopters), initially one and only one of the following labels can be associated to each agent:

- **Perfectly suitable for EV adoption**, if the daily trips of the agent never exceed 200 km;

- **Outliers**, namely agents with homes that are distant from the average house position. Let \( \mu_h = (\text{long}_{\mu_h}, \text{lat}_{\mu_h}) \) be the point whose GPS coordinates are given by
  \[
  \text{long}_{\mu_h} = \frac{1}{|\mathcal{V}|} \sum_{v \in \mathcal{V}} \text{long}_v, \quad \text{lat}_{\mu_h} = \frac{1}{|\mathcal{V}|} \sum_{v \in \mathcal{V}} \text{lat}_v,
  \]
  and denote with \( \sigma_h \) the standard deviation of the distances between the agents’ homes and \( \mu_h \). In this work, we classify a vehicle \( v \in \mathcal{V} \) as outlier if
  \[
  d(h_v, \mu_h) > \sigma_h \text{ [km]}.
  \]

- **Almost isolated agents**, i.e., the nodes such that
  \[
  |\mathcal{N}_v| < 8, \quad v \in \mathcal{V}.
  \]

Note that, by trimming almost isolated nodes, we implicitly assume that agents with few neighbors are unlikely to change their opinion on EV adoption. As a result of this pruning procedure, we obtain the connected graph shown in Figure 1, which comprises 894 vehicles (i.e., approximately 89% of the agents available after preprocessing). The associated distribution of the connection degree is shown in Figure 2, showing that most of the nodes have a relatively low degree, but there is also a fair number of agents that are regarded as proximal to 38% of the overall nodes.
not immediately suitable EV adoption, if the daily trips of the agent exceed 200 km, independently of the possibility of recharge linked to the individual motion patterns.

In this paper, these labels are exploited as a proxy for the personal inclination towards EVs and their adoption. Indeed, it seems likely that people able to perform their daily trips without facing the problems related to the limitations of EVs would be more inclined to adopt this new technology. As a consequence, in principle $S_0$ is constituted by all agents that, at time $t = 0$, are associated to the first label. By categorizing the nodes, the graph in Figure 1 becomes the one reported in Figure 3(b) which shows how initial EV-adopting agents are positioned within the network. Additional insights on the dimension of the seed set is provided in Figure 3(a), which shows that 29.5% of vehicles belong to $S_0$, i.e., 264 out of 894 nodes of the data-driven network.

IV. DATA-DRIVEN NETWORK ANALYSIS

In this section we consider a few scenarios of interest to evaluate the sensitivity of the EV diffusion model described in (1) to different parameters. For our experiments to be consistent with the actual percentage of EVs registration in Italy (namely, 7% of the overall vehicle registrations), $S_0$ is chosen uniformly at random among the drivers that have been classified as perfectly suitable for EV adoption, instead of making the seed set coincide with the cluster $\mathcal{C}_1$ of vehicles classified as perfectly suitable for EV adoption. Specifically, we fix $|S_0| = \eta|\mathcal{C}_1|$, with $\eta = 0.25$, and we perform 20 tests by randomly re-sampling $\mathcal{C}_1$ so as to consider different initial configurations. A specific instance of the network at $t = 0$ is shown in Figure 4(a) that clearly shows the limited number of seeds with respect to the overall number of agents. Initially, we set the parameters $\alpha_v$ to be all equal to $\alpha = 0.15$ for all $v \in \mathcal{V} \setminus S_0$. Figure 5 shows the average percentage $\Delta a [%]$ of EV-adopters (averaged over 20 instances) as a function of time $t \in [0, 0.5]$. In our experiments, each time step corresponds to six months, so that we evaluate EV adoption over a span of around five years, which seems a reasonable assumption considering that technological changes on longer time spans might change the agents perception on EV adoption. Despite the relatively low number of seeds, the overall percentage of EV-adopters increases over time up to 40%.

In a second experiment, we analyze the sensitivity of the model to changes in the threshold value. Specifically, we consider increasing values of $\alpha_v$ within the interval $\alpha_v \in [0.01, 0.7]$, so as to evaluate both settings in which the agents are keen on adopting EVs, i.e., when $\alpha_v = 0.1$ quite few neighbors (10%) must have adopted EVs for an agent to adopt an electric vehicle itself, to situations in which the agents are generally very skeptical about EVs, namely when $\alpha_v = 0.7$. The average percentage of EV diffusion (over 20 instances) after 5 years is shown in Figure 6 as a function of the threshold. It is clear that the more the agent are not convinced by electric vehicles, the less the final percentage of adopters is. It is interesting to note that the values of $\alpha_v$ in the range $\alpha_v \in [0.15, 0.4]$ seem to be critical for EV adoption. Indeed, for $\alpha_v \in (0.15, 0.4)$ the percentage of final adopters sensibly decreases and, for $\alpha_v \geq 0.4$ the opinion does not generally diffuses throughout the network.

As a final test, we consider how EV diffusion is affected by different choices of $\eta$, namely the number of initial EV-adopters, when $\alpha_v = 0.15$. Note that, for $\eta = 0.25$, the choice of this threshold corresponds to a setting in which we get close to reaching consensus over the network within the considered steps (see Figure 6). The final percentage of EV-adopters obtained for $\eta = [0.1, 0.95]$ is reported in Figure 7. It is clear that it is sufficient to have approximately 15% of the perfectly suitable agents as seeds to trigger a cascade adoption over the network.

By looking at Figures 5-7 the expected percentage of diffusion is concentrated around the average value for different random simulation. By considering the initial instance reported in Figure 4(a) which is obtained by fixing $\eta =
0.25, in Figure 4 we show an instance of the obtained opinion diffusion process. Note that, in this specific case study, we still consider a relatively low number of initial EV-adopters, as this reflects the quite realistic scenario in which even drivers that are perfectly suited for EVs as far as driving habits are concerned, do not adopt EVs also due to the absence of social influences. As it can be seen, although few drivers have adopted EVs, after almost 5 years of evolution, the EV adoption diffuses to quite a large portion of the network, while a relatively small set of agents remains resistant to the change. Note that the agents not adopting EVs are those belonging to regions with a very low number of initial seeds, which seems to be too isolated to actually influence their surroundings.

V. INCENTIVE POLICY

Based on the analysis on the sensitivity to parametric changes in the model, we now consider how to assess the effect of providing incentives on the spread of EV-adoption. Our main goal here is to identify, if possible, a set of influential nodes in the network that can guide the choices of the other agents, thus boosting EVs acceptance. Note that, although in this preliminary analysis we study the effect of incentives empirically, this problem is strictly related to the influence maximization problem over cascading networks [14], that can be summarized as follows.
random among the drivers that are perfectly suitable for EV adoption, so that \( |\Gamma|/|\mathcal{V} - S_0| = 0.15 \) (for \( \eta = 0.1 \) and \( \alpha = 0.15 \)).

The final adoption percentage after 5 years is reported in Figure 8(a) as a function of \( \rho \), with the thick curve representing the average percentage of adoption after 5 years (over 100 instances). This result shows that the considered incentive policy generally leads to a relatively small improvement in the adoption of electric vehicles. To exploit the information that can be drawn from the proximity network, we then find the set \( \Gamma \) among the nodes with the highest degree, which is regarded here as a proxy for the centrality of each node. By looking at Figure 8(b), it is clear that even a simple incentive strategy based on the degree of centrality [15] is able to guarantee a large diffusion. The result of this study shows that the targeting strategy plays a key role to boost the adoption process of EVs. Indeed, while the percentage of adoption is limited to the range [0, 0.4] if incentives are given to randomly chosen agents, the use of degree in choosing the agents to be supported allows to have a large gain, leading to final 0.8 percentage of diffusion, with fixed network topology.

VI. CONCLUDING REMARKS

In this paper, a network-based modelling of mobility patterns has been presented, that allows us to model the proximity among different users within a graph-based framework. Based on this, the social network of interest was built based on the description of real mobility data measured from telematic e-boxes, considering how close the estimate home locations of the different users are. This network was then used to setup an analysis of the adoption an electric vehicles from users, starting from the initial suitability to electric mobility that was extracted from the initial data. The opinion diffusion patterns have been analysed under different scenarios, and an incentive policy has been outlined to foster the adoption rate. Ongoing work is being devoted to further investigate the potential of the adopted framework, considering other proximity measures and using also social variables, such as, for example, the users’ income, to refine the adoption mechanism.

REFERENCES

[1] I. Docherty, G. Marsden, and J. Anable, “The governance of smart mobility,” Transportation Research Part A: Policy and Practice, vol. 115, pp. 114–125, 2018.
[2] F. Grelier, CO2 EMISSIONS FROM CARS: the facts, 2018.
[3] U.S. Department of Energy - Office of Energy Efficiency and Renewable energy, Electric-Drive Vehicles, 2017, https://afdc.energy.gov/files/publication/electric_vehicles.pdf.
[4] Z. A. Needell, J. McNerney, M. T. Chang, and J. E. Trancik, “Potential for widespread electrification of personal vehicle travel in the united states,” Nature Energy, vol. 1, no. 9, p. 16112, 2016.
[5] S. Barr and J. Prillwitz, “A smarter choice? exploring the behaviour change agenda for environmentally sustainable mobility,” Environment and Planning C: government and policy, vol. 32, no. 1, pp. 1–19, 2014.
[6] International Energy Agency, Global EV Outlook 2019 - Scaling-up the transition to electric mobility, 2019.
[7] McKinsey & Company, Electrifying insights: How automakers can drive electrified vehicle sales and profitability, 2017.
[8] D. Acemoglu, A. Ozdaglar, and E. Yildiz, “Diffusion of innovations in social networks,” in 2011 50th IEEE Conference on Decision and Control and European Control Conference, 2011.
[9] P. Frasca, C. Ravazzi, R. Tempo, and H. Ishii, “Gossips and prejudices: Ergodic randomized dynamics in social networks,” IFAC Proceedings Volumes, vol. 46, no. 27, pp. 212–219, 2013.

[10] N. E. Friedkin, A. V. Proskurnikov, R. Tempo, and S. E. Parsegov, “Network science on belief system dynamics under logic constraints,” Science, vol. 354, no. 6310, pp. 321–326, 2016.

[11] J. Lorenz, “Continuous opinion dynamics under bounded confidence: A survey,” International Journal of Modern Physics C, vol. 18, no. 12, pp. 1819–1838, 2007.

[12] D. Rosa and A. Giua, “On the spread of innovation in social networks,” IFAC Proceedings Volumes, vol. 46, no. 27, pp. 322 – 327, 2013, 4th IFAC Workshop on Distributed Estimation and Control in Networked Systems (2013).

[13] Z. F., S. S., T. M., F. S., G., and S. M., “Mining the electrification potential of fuel-based vehicles mobility patterns: a data-based approach,” in First IEEE Conference on Human-Machine Systems, 2020, submitted.

[14] Y. Lim, A. Ozdaglar, and A. Teytelboym, “A simple model of cascades in networks,” submitted, 2016.

[15] M. Newman, Networks: An Introduction. New York, NY, USA: Oxford University Press, Inc., 2010.
This figure "fig1.png" is available in "png" format from:

http://arxiv.org/ps/2001.09704v1