Thesis Projects

Szakdolgozat Témák

2025-2026

Contents

OPEN Open Projects / Elérhető témák

EN: AI-Assisted Post-Quantum Cryptography Migration Advisor

HU: MI-alapú tanácsadó rendszer posztkvantum kriptográfiai migrációhoz

EN

Level: BSc TDK / MSc

Why this matters: Most of today's internet security (HTTPS, VPNs, banking) relies on mathematical problems that classical computers cannot solve efficiently, such as factoring very large numbers. Quantum computers will eventually break these protections. To prepare, NIST published its first post-quantum cryptography (PQC) standards in August 2024: new algorithms (ML-KEM, ML-DSA, SLH-DSA) designed to resist quantum attacks. The problem is that migrating to these standards is not a simple software update: different PQC algorithms make different trade-offs between CPU usage, memory, key sizes, and latency, and the right choice depends on the system's hardware and threat profile.

What the student will build: A decision-support tool that takes system constraints (platform type, bandwidth limits, latency requirements, threat assumptions) as input and recommends suitable PQC primitives with predicted performance characteristics. The student will collect benchmark measurements of PQC implementations across multiple platforms (desktop, ARM single-board computer, and optionally a microcontroller), build a dataset, and train a lightweight predictor or rule-based recommender on top of it.

Goals:

- Benchmark PQC algorithms (ML-KEM, ML-DSA, SLH-DSA) across at least two hardware platforms, measuring CPU time, RAM, key/ciphertext sizes, and round-trip latency.
- Build a structured dataset of these measurements with tagged platform and configuration metadata.
- Develop a recommender (ML model or constraint-based) that, given a system profile, outputs a ranked list of safe PQC configurations.
- Validate recommendations against known performance profiles from reference implementations.

HU

Szint: BSc TDK / MSc

Miért fontos: Az internet biztonságának (HTTPS, VPN, banki rendszerek) nagy része olyan matematikai problémákon alapul, amelyeket klasszikus számítógépek nem tudnak hatékonyan megoldani, például nagy számok faktorizálásán. A kvantumszámítógépek azonban képesek lesznek feltörni ezeket a védelmeket. Felkészülésként a NIST 2024 augusztusában publikálta az első posztkvantum kriptográfiai (PQC) szabványait: ML-KEM, ML-DSA és SLH-DSA. A gond az, hogy az átállás nem egyszerű frissítés: a különböző PQC algoritmusok más-más kompromisszumot kínálnak CPU-használat, memória, kulcsméretek és késleltetés terén, és a helyes választás a rendszer hardverétől és fenyegetettségi profiljától függ.

Mit épít a hallgató: Egy döntéstámogató eszközt, amely rendszer-korlátok (platform, sávszélesség, késleltetés, fenyegetettségi feltételezések) alapján ajánl PQC primitíveket, becsült teljesítményjellemzőkkel. A hallgató méréseket végez PQC implementációkkal több platformon, adathalmazt épít, és erre ráültetve egy könnyű előrejelzőt vagy szabályalapú ajánlórendszert készít.

Célok:

- PQC algoritmusok (ML-KEM, ML-DSA, SLH-DSA) benchmarkolása legalább két platformon (CPU-idő, RAM, kulcs/titkosszöveg méret, késleltetés).
- Strukturált mérési adathalmaz felépítése platformcímkékkel és konfigurációs metaadatokkal.
- Ajánlórendszer (ML modell vagy szabályalapú) fejlesztése, amely a rendszerprofil alapján rangsorolt PQC konfigurációket javasol.
- Az ajánlások validálása referencia-implementációk teljesítményprofiljai alapján.

EN: RL-Guided API Security Testing: Fuzzing as a Credit Assignment Problem

HU: RL-vezérelt API biztonsági tesztelés: Fuzzing mint kredit-hozzárendelési probléma

EN

Level: MSc / MSc TDK

Why this matters: Fuzzing is a software testing technique where a program is fed large amounts of random or semi-random input to discover crashes, hangs, or security vulnerabilities. Traditional fuzzers use coverage-guided heuristics (e.g., AFL), but they struggle with structured inputs like REST API call sequences, where the order and parameters of requests matter enormously. This is essentially a sequential decision-making problem, exactly the kind of task reinforcement learning (RL) is designed for. The twist is that the reward signal (a crash or a new code path) is sparse and delayed: the fuzzer may need to send hundreds of requests before anything interesting happens. In RL terminology, this is a credit assignment problem under sparse, delayed feedback.

What the student will build: An RL-guided fuzzer for REST APIs. The agent learns which sequences of API calls, parameter choices, and mutation strategies are most likely to trigger bugs. The student will reproduce a baseline (e.g., coverage-guided random fuzzing), then build an RL agent (DQN or PPO) that treats each API request as an action, coverage/crash events as reward, and compares sample efficiency and bug-finding rate.

Goals:

- Implement a baseline random/coverage-guided REST API fuzzer.
- Design an RL formulation: state = API response/coverage summary, action = next call + parameters, reward = new coverage or crash.
- Train a DQN or PPO agent and compare bug-finding rate and coverage growth versus the baseline.
- Analyse the credit assignment challenge: which request in a sequence actually caused the crash, and does the RL agent learn to identify it?

HU

Szint: MSc / MSc TDK

Miért fontos: A fuzzing egy szoftvertesztelési technika, amely során a programot nagy mennyiségű véletlen vagy félig véletlen bemenettel tápláljuk, hogy összeomlásokat, lefagyásokat vagy biztonsági réseket találjunk. A hagyományos fuzzerek lefedettségvezérelt heurisztikákat használnak (pl. AFL), de nehezen boldogulnak strukturált bemenetekkel, mint a REST API hívássorozatok, ahol a kérések sorrendje és paraméterei döntőek. Ez lényegében szekvenciális döntéshozatali feladat, pont az, amire a megerősítéses tanulás (RL) tervezve van. A nehézség az, hogy a jutalom (egy összeomlás vagy új kódútvonal) ritka és késleltetett: a fuzzernek akár több száz kérést kell küldenie, mielőtt bármi érdekes történik. RL terminológiában ez egy kredit-hozzárendelési probléma ritka, késleltetett visszajelzés mellett.

Mit épít a hallgató: Egy RL-vezérelt fuzzert REST API-khoz. Az ágens megtanulja, hogy mely API-hívás sorozatok, paraméterválasztások és mutációs stratégiák váltják ki legvalószínűbben a hibákat. A hallgató először egy baseline fuzzert reprodukál, majd épít egy RL ágenst (DQN vagy PPO), amely összehasonlítja a hatékonyságot és a hibatalálati rátát.

Célok:

- Baseline random/lefedettségvezérelt REST API fuzzer implementálása.
- RL formalizáció tervezése: állapot = API válasz/lefedettség, akció = következő hívás + paraméterek, jutalom = új lefedettség vagy összeomlás.
- DQN vagy PPO ágens betanítása, és a hibatalálati ráta, illetve lefedettségnövekedés összehasonlítása a baseline-nal.
- A kredit-hozzárendelési kihívás elemzése: a sorozat melyik kérése okozta ténylegesen az összeomlást, és az RL ágens megtanulja-e ezt azonosítani?

EN: Attack-Based Machine Unlearning Verification Benchmark

HU: Támadásalapú gépi unlearning-verifikációs benchmark

EN

Level: MSc / MSc TDK

Why this matters: Machine unlearning promises to "erase" specific data or behaviours from a trained model. But how do you verify that the model truly forgot? Current practice often relies on indirect metrics (accuracy on retain/forget sets), which can miss subtle information leakage. This project treats privacy attacks as unit tests for forgetting: after unlearning, membership inference and model inversion probes are run systematically to measure whether "forgotten" information remains extractable.

What the student will build: A reproducible evaluation harness that (1) trains a model, (2) applies one or more unlearning methods (exact retraining, fine-tuning, gradient ascent, SCRUB-style approaches), and (3) runs a battery of attack-based verification tests. The benchmark reports standardised metrics: forget quality, retain quality, and leak reduction as measured by membership inference accuracy and inversion reconstruction error.

Goals:

- Implement at least three unlearning strategies under a unified API (retraining, fine-tuning, gradient-ascent).
- Build a verification suite: membership inference attack (shadow model approach) + model inversion probe.
- Define and compute standardised metrics: forget quality, retain quality, leak reduction ratio.
- Compare methods on at least two datasets (tabular + image) and publish reproducible results with open-source code.

HU

Szint: MSc / MSc TDK

Miért fontos: A gépi unlearning azt ígéri, hogy "törli" a modellből a megjelölt adatokat vagy viselkedéseket. De hogyan ellenőrizzük, hogy a modell valóban elfelejtett? A jelenlegi gyakorlat gyakran közvetett metrikákra épít (pontosság a megtartandó/elfelejtendő halmazokon), ami nem mindig veszi észre a finom információszivárgást. A projekt adatvédelmi támadásokat használ a felejtés egységtesztjeiként: unlearning után szisztematikusan lefuttatott membership inference és model inversion próbákkal méri, hogy a „törölt" információ kinyerhető-e még.

Mit épít a hallgató: Egy reprodukálható kiértékelő keretrendszert, amely (1) betanít egy modellt, (2) alkalmaz egy vagy több unlearning módszert (újratanítás, fine-tuning, gradiens-felszállás, SCRUB-szerű megközelítések), majd (3) lefuttat egy támadásalapú verifikációs tesztcsomagot. A benchmark szabványosított metrikákat ad: felejtésminőség, megtartás-minőség és szivárgáscsökkentés membership inference pontosság és inverzió rekonstrukciós hiba alapján.

Célok:

- Legalább három unlearning stratégia implementálása egységes API alatt (újratanítás, fine-tuning, gradiens-felszállás).
- Verifikációs csomag: membership inference támadás (shadow model) + model inversion próba.
- Szabványosított metrikák definiálása és kiszámítása: felejtésminőség, megtartás-minőség, szivárgáscsökkentési arány.
- Módszerek összehasonlítása legalább két adathalmazon (táblás + kép), reprodukálható eredményekkel és nyílt forráskóddal.

EN: Procedural Game Level Generation with Reinforcement Learning

HU: Procedurális játékpálya-generálás megerősítéses tanulással

EN

Level: BSc TDK / MSc

Why this matters: Game levels are usually designed by hand, which is slow and expensive. Procedural content generation (PCG) can automate this, but naive random generation often produces unplayable or boring results. The idea is to train an RL agent that builds levels step by step, placing tiles, obstacles, enemies, and objectives, with reward shaped for playability, balance, and variety.

This is also a credit assignment problem: if the final level is unplayable, was it the bad wall at step 5 or the missing exit at step 40? The agent must learn to attribute outcomes to specific design decisions, a challenge that connects directly to reward shaping and temporal credit assignment research in RL.

What the student will build: A pipeline where a PPO or DQN agent sequentially constructs 2D grid-based game levels. Playability is checked by an automated solver (A* or BFS pathfinding). The agent receives shaped rewards for reachability, difficulty balance, and aesthetic diversity. The student will compare the RL-generated levels to pure random and search-based baselines (e.g., evolutionary generation).

Goals:

- Design a grid-world level representation and a step-by-step placement action space.
- Implement automated playability verification (pathfinding from start to goal).
- Train an RL agent (PPO or DQN) that generates playable, diverse levels.
- Compare against random generation and an evolutionary/search-based baseline on playability rate, diversity, and difficulty distribution.
- Analyse credit assignment: which placement steps does the agent learn are most critical for level quality?

HU

Szint: BSc TDK / MSc

Miért fontos: A játékpályákat általában kézzel tervezik, ami lassú és költséges. A procedurális tartalomgenerálás (PCG) automatizálhatja a folyamatot, de a naiv véletlenszerű generálás gyakran játszhatatlan vagy unalmas eredményt ad. A cél egy olyan RL ágens tanítása, amely lépésről lépésre építi a pályát: csempéket, akadályokat, ellenségeket és célokat helyez el, és jutalmat kap, ha az eredmény játszható, kiegyensúlyozott és érdekes. Ez egyben kredit-hozzárendelési probléma is: ha a végső pálya játszhatatlan, az az 5. lépésben rossz helyre tett fal miatt van, vagy a 40. lépésben hiányzó kijárat miatt? Az ágensnek meg kell tanulnia az eredményt konkrét tervezési döntésekhez kapcsolni, ami közvetlenül kötődik a jutalomalakítás és az időbeli kredit-hozzárendelés kutatásához.

Mit épít a hallgató: Egy pipeline-t, amelyben egy PPO vagy DQN ágens lépésről lépésre épít 2D rácsalapú játékpályákat. A játszhatóságot automatikus megoldó ellenőrzi (A* vagy BFS útkereséssel). Az ágens formált jutalmat kap elérhetőségért, nehézségi egyensúlyért és változatosságért. A hallgató összehasonlítja az RL-generált pályákat véletlenszerű és keresésalapú (pl. evolúciós) baseline-okkal.

Célok:

- Rácsvilág pálya-reprezentáció és lépésenkénti elhelyezési akciótér tervezése.
- Automatikus játszhatóság-ellenőrzés implementálása (útvonalkeresés starttól a célig).
- RL ágens (PPO vagy DQN) betanítása, amely játszható, változatos pályákat generál.
- Összehasonlítás véletlenszerű generálással és evolúciós/keresésalapú baseline-nal játszhatósági ráta, változatosság és nehézségeloszlás mentén.
- Kredit-hozzárendelés elemzése: mely elhelyezési lépéseket tanulja meg az ágens kritikusnak a pálya minőségéhez?

EN: Shapley-Based Credit Assignment for Cooperative Multi-Agent Teams

HU: Shapley-alapú kredit-hozzárendelés kooperatív többügynökös csapatokhoz

EN

Level: MSc / MSc TDK

Why this matters: In cooperative multi-agent reinforcement learning (MARL), the entire team often shares a single global reward. But each agent needs to know how much it contributed. When this signal is noisy or misleading, agents may become "lazy" (free-riding on teammates) or learn unstable policies. One principled solution comes from cooperative game theory: Shapley values, which compute each agent's average marginal contribution across all possible coalitions. The problem is that the exact Shapley value is exponentially expensive to compute, so practical systems use sampling-based approximations.

What the student will build: A sampling-based Shapley credit estimator integrated into a MARL training loop (PPO or QMIX backbone). The estimator replaces or augments the global reward with per-agent Shapley-based credits. The student will measure whether this improves sample efficiency, training stability, and role specialisation compared to: (i) naive global reward, (ii) COMA-style counterfactual baselines, and (iii) QMIX value factorisation. Experiments will run on PettingZoo MPE cooperative tasks and at least one SMAC micromanagement map.

Goals:

- Implement a sampling-based Shapley value estimator with configurable coalition sample budgets.
- Integrate Shapley credits into PPO or QMIX training as per-agent shaped rewards.
- Compare against global reward, COMA, and QMIX on at least two cooperative benchmarks (PettingZoo + SMAC).
- Analyse role specialisation emergence, lazy-agent frequency, and runtime-quality trade-offs from different sampling budgets.

HU

Szint: MSc / MSc TDK

Miért fontos: A kooperatív többügynökös megerősítéses tanulásban (MARL) a csapat gyakran egyetlen globális jutalmat kap. De minden ágensnek tudnia kell, mennyit járult hozzá ő. Ha ez a jel zajos vagy félrevezető, az ágensek „lusták" lehetnek (potyautasok), vagy instabil policy-kat tanulhatnak. Egy elvi megoldás a kooperatív játékelméletből származik: a Shapley-értékek, amelyek minden ágens átlagos marginális hozzájárulását számolják ki az összes lehetséges koalíción keresztül. A probléma az, hogy a pontos Shapley-érték kiszámolása exponenciálisan költséges, ezért a gyakorlati rendszerek mintavétel-alapú közelítéseket használnak.

Mit épít a hallgató: Egy mintavétel-alapú Shapley kredit-becslőt, amely egy MARL tanítási ciklusba (PPO vagy QMIX gerinc) van integrálva. A becslő az ágensenkénti Shapley-kreditekkel helyettesíti vagy kiegészíti a globális jutalmat. A hallgató méri, hogy ez javítja-e a mintahatékonyságot, a tanítási stabilitást és a szerepspecializációt a következőkhöz képest: (i) naiv globális jutalom, (ii) COMA-stílusú kontrafaktuális baseline-ok, (iii) QMIX értékfaktorizáció. A kísérleteket PettingZoo MPE kooperatív feladatokon és legalább egy SMAC térképen futtatjuk.

Célok:

- Mintavétel-alapú Shapley-érték becslő implementálása konfigurálható koalíciós mintaszámmal.
- Shapley-kreditek integrálása PPO vagy QMIX tanításba ágensenkénti formált jutalomként.
- Összehasonlítás globális jutalommal, COMA-val és QMIX-szel legalább két kooperatív benchmarkon (PettingZoo + SMAC).
- Szerepspecializáció-megjelenés, potyautaság gyakorisága, és futásidő-minőség kompromisszumok elemzése különböző mintaszámokkal.

EN: Adversarial Robustness of ML-Based Malware Detectors

HU: ML-alapú malware-detektorok ellenállósága adverzariális támadásokkal szemben

EN

Level: BSc TDK / MSc

Why this matters: Modern malware detection increasingly relies on ML classifiers trained on features extracted from executables (byte sequences, API call graphs, PE header fields), not just hand-written signatures. These classifiers reach high accuracy on benchmarks. But are they robust? Adversarial ML shows that small, carefully chosen input modifications can flip a classifier's decision while preserving the input's real functionality. In the malware context, an attacker could take a known-malicious binary, apply adversarial perturbations (appending bytes, reordering sections, inserting no-op API calls), and make the detector classify it as clean.

What the student will build: A three-stage pipeline: (1) train a baseline ML malware classifier on a public dataset (e.g., EMBER, BODMAS, or MalBench), (2) implement at least two evasion attack strategies (gradient-based perturbation on feature space + a practical "functionality-preserving" mutation like section appending or import table padding), and (3) evaluate and compare defences (adversarial training, input preprocessing/feature squeezing, ensemble voting).

Goals:

- Train a malware classifier (random forest + a small neural network) on at least one public malware feature dataset.
- Implement two evasion attacks: a white-box gradient-based attack on the feature vector, and a black-box functionality-preserving binary mutation.
- Measure detection rate degradation under each attack (clean accuracy vs adversarial accuracy).
- Implement and evaluate at least two defences (adversarial training, feature squeezing or ensemble), reporting recovered accuracy.

HU

Szint: BSc TDK / MSc

Miért fontos: A modern malware-detektálás egyre inkább ML osztályozókra támaszkodik, amelyeket futtatható fájlokból kinyert jellemzőkön (bájt-szekvenciák, API-hívás gráfok, PE fejlécmezők) tanítanak, nem csak kézzel írt szignatúrákra. Ezek az osztályozók magas pontosságot érnek el a benchmarkokon. De vajon robosztusak-e? Az adverzariális gépi tanulás megmutatja, hogy kis, gondosan megválasztott módosításokkal át lehet billenteni az osztályozó döntését, miközben a bemenet valódi funkcionalitása megmarad. Malware-kontextusban egy támadó egy ismert kártevőt úgy módosíthat (bájtok hozzáfűzése, szekciók átrendezése, no-op API-hívások beszúrása), hogy a detektor "tisztának" minősítse.

Mit épít a hallgató: Egy háromfázisú pipeline-t: (1) baseline ML malware-osztályozó tanítása publikus adathalmazon (pl. EMBER, BODMAS vagy MalBench), (2) legalább két kitérő (evasion) támadás implementálása (gradiens-alapú perturbáció a jellemzőtéren + gyakorlati, funkcionalitásmegőrző mutáció, mint szekció-bővítés vagy importtábla-kitöltés), (3) védelmek kiértékelése és összehasonlítása (adverzariális tanítás, bemenet-előfeldolgozás/feature squeezing, ensemble szavazás).

Célok:

- Malware-osztályozó tanítása (random forest + kis neurális hálózat) legalább egy publikus malware jellemzőadathalmazon.
- Két kitérő támadás implementálása: white-box gradiens-alapú támadás a jellemzővektoron, és black-box funkcionalitásmegőrző bináris mutáció.
- Detekciós ráta csökkenésének mérése minden támadás alatt (tiszta pontosság vs. adverzariális pontosság).
- Legalább két védelmi módszer implementálása és kiértékelése (adverzariális tanítás, feature squeezing vagy ensemble), a visszanyert pontosság riportolásával.

EN: Detecting and Classifying LLM Prompt Injection Attacks

HU: LLM prompt injection támadások detektálása és osztályozása

EN

Level: BSc / BSc TDK

Why this matters: Large language models (LLMs) like ChatGPT, Gemini, and Claude are being embedded into web applications, customer support bots, coding assistants, and enterprise tools at a rapid pace. A new class of security vulnerability has emerged with them: prompt injection. The idea is simple but dangerous. An attacker crafts input text that "hijacks" the model's instructions, making it ignore its original system prompt and instead follow the attacker's commands. This can leak confidential context, bypass content filters, or cause the model to take unauthorised actions (e.g., calling APIs it shouldn't). OWASP now lists prompt injection as the #1 risk for LLM applications.

What the student will build: A prompt injection detection and classification system. The student will (1) curate a labelled dataset of prompt injection attempts (from public collections like Garak, HackAPrompt challenge datasets, Prompt Injection Benchmark, and manually crafted samples covering direct injection, indirect injection, jailbreaks, and payload-smuggling), (2) train and compare detection models (fine-tuned text classifiers: DistilBERT or a small LLM adapter, plus simpler baselines like TF-IDF + logistic regression and perplexity-based filters).

Goals:

- Curate a labelled dataset (≥2000 samples) covering at least four prompt injection categories: direct injection, indirect/context injection, jailbreak, and payload smuggling.
- Train baseline detectors (TF-IDF + logistic regression, perplexity filter) and a fine-tuned transformer classifier (DistilBERT or similar).
- Report precision, recall, F1, and false positive rate, and compare all approaches on a held-out test set.

HU

Szint: BSc / BSc TDK

Miért fontos: A nagy nyelvi modelleket (LLM-ek), mint a ChatGPT, Gemini vagy Claude, egyre gyorsabban integrálják webalkalmazásokba, ügyfélszolgálati chatbotokba, kódolási asszisztensekbe és vállalati eszközökbe. Ezzel egy új típusú biztonsági sérülékenység is megjelent: a prompt injection. A lényeg egyszerű, de veszélyes. A támadó úgy fogalmazza meg a bemenetét, hogy az „eltéríti" a modell utasításait: figyelmen kívül hagyatja az eredeti rendszerpromptot, és helyette a támadó parancsait követi. Ezzel bizalmas kontextus szivároghat ki, tartalomszűrők kerülhetők meg, vagy a modell jogosulatlan műveleteket hajthat végre (pl. API-hívások). Az OWASP a prompt injectiont jelenleg az LLM-alkalmazások #1 kockázataként tartja számon.

Mit épít a hallgató: Prompt injection detektáló és osztályozó rendszert. A hallgató (1) címkézett adathalmazt állít össze prompt injection kísérletekből (publikus gyűjteményekből, mint a Garak, HackAPrompt, Prompt Injection Benchmark, valamint kézzel készített mintákból, lefedve a direkt injekciót, indirekt injekciót, jailbreaket és payload-csempészést), (2) detekciós modelleket tanít és hasonlít össze (fine-tuned szövegosztályozó: DistilBERT vagy kis LLM adapter, plusz egyszerűbb baseline-ok: TF-IDF + logisztikus regresszió, perplexitásalapú szűrő).

Célok:

- Címkézett adathalmaz összeállítása (≥2000 minta) legalább négy prompt injection kategóriában: direkt injekció, indirekt/kontextus injekció, jailbreak, payload-csempészés.
- Baseline detektorok tanítása (TF-IDF + logisztikus regresszió, perplexitásszűrő) és fine-tuned transformer osztályozó (DistilBERT vagy hasonló).
- Precízió, recall, F1 és hamis pozitív ráta riportolása, valamint az összes megközelítés összehasonlítása tartott teszthalmazon.

EN: Neural Distinguishers for Lightweight Ciphers

HU: Neurális megkülönböztetők könnyűsúlyú rejtjelezőkhöz

EN

Level: BSc TDK / MSc
Why this matters: A good cipher's output should be indistinguishable from random noise. In 2019 Gohr showed that a neural network can learn to tell apart real SPECK ciphertext from random data, a task classical statistical tests fail at for reduced-round variants. This "neural distinguisher" approach has since been applied to other lightweight ciphers (SIMON, PRESENT, GIFT) and is one of the hottest directions in AI-assisted cryptanalysis. The student trains binary classifiers (MLP, CNN, ResNet-style) that receive a ciphertext pair and predict "real cipher output" vs "random". No physical hardware is needed. All data is generated in software.

Goals:

- Generate ciphertext/random datasets for at least two lightweight ciphers (e.g., SPECK-32/64 and SIMON-32/64) at varying round counts.
- Train neural distinguishers and report accuracy vs round count (find the "break point" where the network can no longer distinguish).
- Compare against classical statistical distinguishers (e.g., χ^2 test, linear approximation).
- Analyse what the network learns via gradient-based input attribution (which ciphertext bits matter most).

HU

Szint: BSc TDK / MSc
Miért fontos: Egy jó rejtjelező kimenetének megkülönböztethetetlennek kell lennie a véletlentől. 2019-ben Gohr megmutatta, hogy egy neurális hálózat meg tudja különböztetni a valódi SPECK rejtjelszöveget a véletlen adatoktól, olyan feladat, amelyben a klasszikus statisztikai tesztek csődöt mondanak csökkentett körszámú változatoknál. Azóta ezt a „neurális megkülönböztető" megközelítést más könnyűsúlyú rejtjelezőkre is alkalmazták (SIMON, PRESENT, GIFT), és az MI-alapú kriptanalízis egyik legforróbb iránya.
A hallgató bináris osztályozókat tanít (MLP, CNN, ResNet-jellegű), amelyek rejtjelszöveg-párt kapnak, és megjósolják: „valódi rejtjelszöveg" vagy „véletlen". Nincs szükség fizikai hardverre. Minden adat szoftveresen generálható.

Célok:

- Rejtjelszöveg/véletlen adathalmazok generálása legalább két könnyűsúlyú rejtjelezőhöz (pl. SPECK-32/64 és SIMON-32/64) változó körszámmal.
- Neurális megkülönböztetők tanítása és pontosság riportolása körszám függvényében (a „töréspont" meghatározása, ahol a hálózat már nem tud különbséget tenni).
- Összehasonlítás klasszikus statisztikai megkülönböztetőkkel (pl. χ^2 teszt, lineáris approximáció).
- A hálózat tanulásának elemzése gradiens-alapú attribúcióval (mely rejtjelszöveg-bitek számítanak leginkább).

ONGOING Ongoing Projects / Folyamatban lévő szakdolgozatok

EN: Football Management Strategies via MARL

HU: Futballmenedzsment stratégiák tanulása MARL-lel

EN

Name: Baranyi Sándor
Level: MSc Thesis
Topic: This research studies multi-agent reinforcement learning for strategic decision-making in football management. Autonomous “manager agents” learn to choose and adjust tactics in a lightweight simulation inspired by real football data and seasonal dynamics.
Scope and research directions:

- Learning against preset tactical opponents representing fixed managerial styles (stable baselines).
- Adversarial league-style multi-agent training, where multiple managers adapt simultaneously across seasons.
- Seasonality and scheduling variability (match congestion affects fatigue, recovery, and injury risk, requiring long-horizon adaptation across varying calendars).
- Explore between-season adaptation (policy initialization / tactical priors) based on performance and fitness outcomes.

HU

Név: Baranyi Sándor
Szint: MSc szakdolgozat
Téma: A dolgozat többügynökös megerősítéses tanulással (MARL) vizsgálja a futballmenedzsment stratégiai döntéshozatalát. A cél autonóm „menedzser ágensek” tanítása, akik egy könnyű szimulációs környezetben képesek taktikai döntéseket hozni és azokat a körülményekhez igazítani.
Vizsgált irányok / kiterjesztések:

- Rögzített, előre definiált taktikai ellenfelek ellen (stabil baseline-ok, eltérő menedzserstílusok).
- Ligaszerű, többügynökös (adverzariális) beállítás, ahol több menedzser egyszerre tanul és adaptálódik szezonokon át.
- Szezonálisság és menetrend-variabilitás (a mérkőzéssűrűség és a naptártorlódás hat a fáradtságra, regenerációra és sérüléskockázatra, így hosszú távú, szezonon átívelő stratégiaváltást igényel).
- Szezonok közötti adaptáció vizsgálata (policy inicializáció vagy taktikai preferenciák frissítése) teljesítmény- és terhelésmutatók alapján.

EN: Chess Variants with Deep RL in Python

HU: Sakkvariánsok modellezése mély megerősítéses tanulással Python környezetben

EN

Name: Virág Levente
Level: BSc Thesis (2026)
Scope and research directions: Classic chess engines often rely on heuristic search (e.g., minimax, Monte Carlo), which can struggle with long-horizon planning and variant-specific dynamics. This thesis explores deep reinforcement learning to train agents for selected chess variants (including fairy chess), integrated into a playable Pygame framework.
System scope:

- A Pygame-based chess platform supporting multiple variants and match modes (human vs AI, AI vs AI).
- Variant rule implementation (piece sets, move rules, win conditions) in a clean, extensible way.
- RL training pipeline(s) for chosen variants, with evaluation of playing strength and generalization.

HU

Név: Virág Levente
Szint: BSc szakdolgozat (2026)
Téma: A klasszikus sakkprogramok fejlesztéséhez gyakran heurisztikus algoritmusokat (pl. Minimax, Monte Carlo) használnak, amelyek korlátozottak lehetnek hosszabb lépéssorozatok előrejelzésében („horizont-hatás”), és a variánsok sajátos dinamikáját külön kezelést igényelheti. A dolgozat célja mély megerősítéses tanulás alkalmazása különböző sakkvariánsokban (fairy chess), és ezek integrálása egy Pygame alapú játszható keretrendszerbe.
Vizsgált irányok / kiterjesztések:

- Több variánst támogató Pygame sakkplatform, ember-gép és gép-gép módokkal.
- Variáns-szabályok implementálása (bábukészlet, lépésszabályok, nyerési feltételek) bővíthető formában.
- RL tanítási pipeline(ok) kiválasztott variánsokra, majd játékerő és általánosíthatóság értékelése.

FINISHED **Finished Projects / Befejezett szakdolgozatok**

EN: Multi-Agent Board-Game Research Platform (Risk)

HU: Többágenses társasjáték szimulációs és kutatási rendszer: Rizikó elemzése mély RL-lel

EN

Name: Nagy Richárd Bendegúz
Level: BSc Thesis
Topic: Develop a multi-agent simulation and research platform for the classic Risk board game, and use it to analyze strategies with reinforcement learning. The project combines a Python research environment with an interactive web application.
Scope and research directions:

- Research simulator in Python implementing Risk rules and enabling multi-agent matches.
- Multiple agent types: random, rule-based, greedy, and learned (TorchRL encouraged).
- Web application (modern JS frontend) for interactive analysis and human-vs-agent gameplay.
- Not only “find an optimal strategy”, but study game dynamics, emergent strategies, and adaptation between competing agents. Explore hierarchical RL with subgoals (e.g., battles vs. troop movement) for strategic decomposition.

HU

Név: Nagy Richárd Bendegúz
Szint: BSc szakdolgozat
Téma: A szakdolgozat célja egy többágenses társasjáték szimulációs és kutatási rendszer fejlesztése, és ennek felhasználása a klasszikus Rizikó játék stratégiai elemzésére megerősítéses tanulással. A rendszer egy Python nyelvű kutatási szimulátorból és egy interaktív, modern frontenddel készülő webalkalmazásból áll.
Vizsgált irányok / kiterjesztések:

- A Rizikó szabályrendszerének modellezése Pythonban és több ágens párhuzamos futtatása.
- Különböző döntéshozók: véletlenszerű, szabályalapú, mohó és tanult ágensek (TorchRL javasolt).
- Webes UI emberi játékhoz és a kísérletek/eredmények interaktív vizsgálatához.
- Nem pusztán egy „optimális stratégia” keresése, hanem a játék dinamikájának feltárása, emergens stratégiák megjelenésének vizsgálata, valamint az ágensek egymáshoz való adaptációjának elemzése. Hierarchikus RL vizsgálata több alcél és prioritás mentén (pl. csaták, katonák áthelyezése).

EN: Improving NPC Behavior for Immersion via MARL

HU: Nem játékos karakterek viselkedésének fejlesztése MARL-lel az immerzió növeléséért

EN

Name: Tokodi Gergely
Level: MSc Thesis
Topic: Many games rely on deterministic, scripted NPC behaviors that become predictable over long play sessions, reducing immersion. This thesis investigates multi-agent reinforcement learning (MARL) to create NPCs that react to the world and to each other more adaptively in cooperative and competitive tasks.
Scope and research directions:

- Cooperative multi-agent tasks (agents must coordinate to complete objectives rather than looping patrol-like routines).
- Interaction-rich settings (agents influence each other and the environment, producing less repetitive behavior).
- Communication layer to support more believable coordination and interaction.
- Integration of higher-level social/motivation models (e.g., needs/motivation) as an extension.

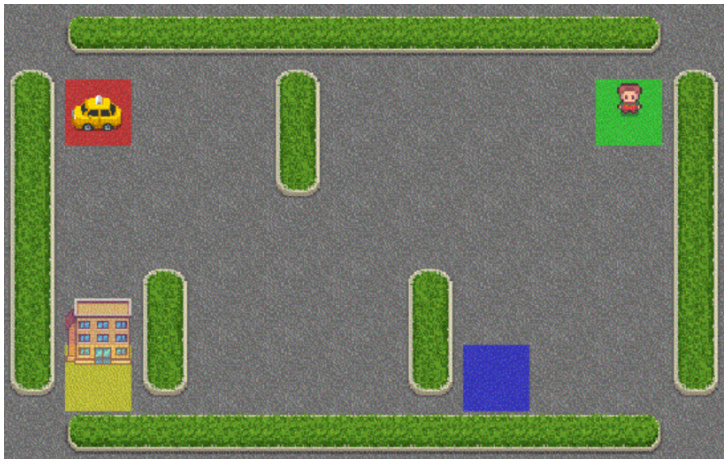
HU

Név: Tokodi Gergely
Szint: MSc szakdolgozat
Téma: A videojátékok NPC-i gyakran determinisztikus, scriptelt viselkedést követnek, ami hosszabb játékidő alatt ismétlődő, kiismerhető mintázatokhoz vezet, és csökkenti az immerziót. A dolgozat célja többügynökös megerősítéses tanulással olyan NPC-viselkedések létrehozása, amelyek adaptívabban és természetesebben reagálnak a világra és egymásra kooperatív és kompetitív helyzetekben.
Vizsgált irányok / kiterjesztések:

- Kooperatív többügynökös feladatok, ahol a szereplőknek együtt kell megoldaniuk egy konkrét célfeladatot.
- Interakció-központú működés (a szereplők egymásra és a környezetre is reagálnak, nem csak ismétlődő rutinokat futtatnak).
- Kommunikáció támogatása, hogy az együttműködés és „beszélgetések” életszerűbbek legyenek.
- (Opcionális) Szociális/motivációs modellek bevonása (pl. Maslow-piramis) a viselkedés finomhangolására.

EN: Targeted Reinforcement Unlearning

HU: Célzott megerősítéses unlearning



EN

Name: Rafig Babayev
Level: MSc TDK
Topic: This thesis develops and evaluates targeted unlearning methods in reinforcement learning to selectively remove specific behaviors or knowledge, motivated by privacy/regulatory needs and changing requirements.
Main directions:

- State-based behavioral unlearning (forget behaviors tied to specific state regions)
- Environment poisoning interventions (controlled environment modifications to induce forgetting)
- Sub-optimal habit removal (remove inefficient learned routines while preserving performance)

Evaluation: Unlearning effectiveness is tested via environment inference / reconstruction-style attacks and behavioral probes across benchmark tasks (e.g., CartPole, MountainCar, LunarLander, Taxi, Acrobot) using DQN/PPO.
Goals:

- Implement a unified experimental framework for targeted RL unlearning.
- Compare multiple unlearning strategies under consistent metrics.

HU

Név: Rafig Babayev
Szint: MSc TDK
Téma: A dolgozat célzott unlearning módszereket fejleszt és hasonlít össze megerősítéses tanulásban, hogy az ágens szelektíven „elfelejtsen” bizonyos viselkedéseket vagy tudást, miközben a megtartandó teljesítmény lehetőleg nem romlik.
Fő irányok:

- Állapothoz kötött viselkedés-elfelejtés (state-space régiókhoz kötött szokások eltávolítása)
- Környezetmódosításos/„poisoning” megközelítések (kontrollált beavatkozás a felejtés kiváltására)
- Szuboptimális szokások eltávolítása (nem hatékony rutinok kiirtása)

Célok:

- Egységes keretrendszer kialakítása célzott RL unlearning kísérletekhez.
- Módszerek összehasonlítása egységes metrikák mentén.

EN: Reward Design Strategies in Cooperative MARL

HU: Jutalom-tervezési stratégiák kooperatív MARL keretrendszerben

EN

Name: Mielec Attila
Level: BSc Thesis (2025)
Topic: An interactive visualization and experimentation platform for multi-agent reinforcement learning, focusing on how reward design and reward shaping affect learning dynamics in cooperative and mixed settings. The system supports PPO-based training and real-time analysis in PettingZoo MPE environments (notably Simple Tag and Simple Spread).
Scope and research directions:

- Build a web-based MARL platform that makes training dynamics observable via live environment rendering, metrics, and episode replay.
- Enable systematic comparison of reward configurations and training hyperparameters, with stored runs for later analysis.
- Demonstrate the approach using PPO and representative multi-agent benchmarks, highlighting the practical impact of reward shaping.

Implementation highlights: Backend in Python (FastAPI) with real-time WebSocket streaming, frontend in React (Tailwind), runs stored via SQLite/JSON.

HU

Név: Mielec Attila
Szint: BSc szakdolgozat (2025)
Téma: Interaktív, vizualizáció-központú kísérleti platform többágenses megerősítéses tanuláshoz, amely azt vizsgálja, hogyan hat a jutalomfüggvény kialakítása és a jutalomalakítás a tanulási dinamikára kooperatív és vegyes szcenáriókban. A rendszer PPO-alapú tanítást és valós idejű elemzést támogat PettingZoo MPE környezetekben (kiemelten: Simple Tag, Simple Spread).
Vizsgált irányok / kiterjesztések:

- Olyan webes MARL platform fejlesztése, ahol a tanulás folyamata követhető élő környezet-megjelenítéssel, metrikákkal és visszajátszással.
- Jutalomkonfigurációk és tréning beállítások összehasonlíthatósága, futások mentése későbbi elemzéshez.
- A jutalomalakítás gyakorlati hatásának bemutatása PPO-val többágenses benchmarkokon.

Technikai váz: Python (FastAPI) backend valós idejű WebSocket kommunikációval, React (Tailwind) frontend, tárolás SQLite/JSON.